

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2003, 14:13:08 ; Search time 41 Seconds  
(without alignments)  
27.100 Million cell updates/sec

Title: US-09-620-840C-2

Perfect score: 17

Sequence: 1 XEXXXLL 7

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

il number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A\_Geneseq\_19Jun03.\*

1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*  
24: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	13	76.5	7	ABU07728	Human leupaxin sec
2	13	76.5	10	AA05815	Human complementar
3	13	76.5	11	ABU08967	Human complement C
4	13	76.5	12	ABG72971	Human complement C
5	13	76.5	14	AA000721	Human protein frag
6	13	76.5	14	AA000722	Human protein frag
7	13	76.5	14	ABP58087	Complement C3f fra
8	13	76.5	15	ABG75803	Congestive heart f
9	13	76.5	17	AA05786	Human calcium bind

10	13	76.5	17	24	ABU08617	Disease specific b
11	13	76.5	17	24	ABB82769	Congestive heart f
12	13	76.5	18	24	ABU08618	Disease specific b
13	13	76.5	19	17	AA057974	Antigenic fragment
14	13	76.5	20	22	AA05792	Human calcium bind
15	13	76.5	25	21	AA05790	Human secreted pro
16	13	76.5	26	22	AA05790	Human secreted pro
17	13	76.5	32	21	AA05790	Human immune/haema
18	13	76.5	34	22	AA05790	Human secreted pep
19	13	76.5	42	13	AA05790	Human polypeptide
20	13	76.5	46	22	AA05790	Peptide sequence e
21	13	76.5	53	22	AA05790	Human liver peptide
22	13	76.5	53	22	AA05790	Peptide #2884 enco
23	13	76.5	53	22	AA05790	Peptide #2902 enco
24	13	76.5	53	22	AA05790	Human nervous syst
25	13	76.5	53	22	AA05790	Protein #2835 enco
26	13	76.5	53	22	AA05790	Human brain expres
27	13	76.5	53	22	AA05790	Human bone marrow
28	13	76.5	53	22	AA05790	Peptide #2842 enco
29	13	76.5	53	22	AA05790	Peptide #2944 enco
30	13	76.5	53	22	AA05790	Peptide #2824 enco
31	13	76.5	64	22	AA05790	Human immune/haema
32	13	76.5	66	21	AA05790	Arabidopsis thalia
33	13	76.5	66	22	AA05790	Proteinibacterium
34	13	76.5	66	22	AA05790	Novel human diagno
35	13	76.5	67	22	AA05790	Human liver peptid
36	13	76.5	67	22	AA05790	Proteinibacterium
37	13	76.5	67	22	AA05790	Peptide #11390 enc
38	13	76.5	67	22	AA05790	Human brain expres
39	13	76.5	67	22	AA05790	Human bone marrow
40	13	76.5	67	22	AA05790	Human immune/haema
41	13	76.5	67	22	AA05790	Peptide #11843 enc
42	13	76.5	67	23	AA05790	Human peptide enco
43	13	76.5	68	22	AA05790	Novel human enzyme
44	13	76.5	69	21	AA05790	Human secreted pro
45	13	76.5	69	22	AA05790	Human polypeptide

#### ALIGNMENTS

#### RESULT 1

ABU07728

ID ABU07728 standard; Peptide; 7 AA.

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

PT New purified and isolated leupaxin polypeptide, useful for identifying  
PT specific binding partner compound of leupaxin that modulates activity  
PT or expression of leupaxin -  
XX Example 12; Page 11; 25pp; English.  
XX The invention relates to a purified and isolated mammalian leupaxin  
XX polypeptide. Leupaxin function is thought to be regulated by tyrosine  
XX kinase activity. Tyrosine phosphorylation of leupaxin is thought to occur  
XX in the induction and process of cell adhesion, motility and growth. The  
XX polypeptide and its polynucleotide are useful for identifying their  
XX specific binding partner compounds. The specific binding partner  
XX modulates activity of the polypeptide or expression of the  
XX polynucleotide. The polypeptide or the polynucleotide are useful in the  
XX systematic analysis of the structure and function of leupaxin, or for  
XX identifying modulators of leupaxin that are therapeutically useful in  
XX treatment of a wide range of diseases and physiological conditions in  
XX which leupaxin biological activity is known to be involved. The  
XX polynucleotide is useful for modifying cells to modulate, increase or  
XX decrease, expression of endogenous leupaxin in host cells which naturally  
XX include polynucleotides that encode leupaxin, for developing through e.g.  
XX homologous recombination or knockout strategies, animals that fail to  
XX express functional leupaxin or that express a variant of leupaxin, for  
XX identifying genomic DNA sequences encoding leupaxin and leupaxin  
XX expression control regulatory sequences such as promoters, operators,  
XX enhancers or repressors, for isolating DNAs encoding allelic variants of  
XX leupaxin, in hybridisation assays to detect the capacity of cells to  
XX express leupaxin, or in diagnostic methods for identifying a genetic  
XX alteration(s) in a leupaxin locus that underlies a disease state or  
XX states. A host cell transformed with the polynucleotide is useful as a  
XX source of immunogen for development of antibodies specifically  
XX immunoreactive with leupaxin, or for large scale production of the  
XX polypeptide. The present sequence represents the amino acid sequence of  
XX mutated human leupaxin second leucine-aspartate (LD) motif which is  
XX used to replace the wild type sequence (see ABU07726).  
XX Sequence 7 AA;

Query Match 76.5%; Score 13; DB 24; Length 7;  
Best Local Similarity 50.0%; Pred. No. 9.3e+05;  
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7  
Db 2 EAAALL 7

RESULT 2  
95815  
AAG95815 standard; Peptide; 10 AA.

XX AAG95815;  
XX 18-SEP-2001 (first entry)  
XX Human complementary peptide, SEQ ID NO: 2009.  
XX Human; complementary peptide; ligand; drug discovery; drug design.  
XX Homo sapiens.  
XX WO200142277-A2.  
XX 14-JUN-2001.  
XX 13-DEC-2000; 2000WO-GB04776.  
XX 13-DEC-1999; 99GB-0029464.  
XX (PROT-) PROTEOM LTD.  
XX Roberts GW, Heal JR;

WPI; 2001-408419/43.

XX A set of peptide ligands consisting of specific complementary peptides  
PT to proteins encoded by genes of the human genome, useful in an assay  
PT for screening and identifying of one or more novel peptides which are  
XX drug candidates or pro-drugs -

XX Example 4; Page 330; 646pp; English.

XX The invention relates to a set of complementary peptide ligands  
XX generated from the human genome. The complementary peptides  
XX interact with their relevant target proteins encoded in the human  
XX genome. They can be used as reagents in drug discovery and as lead  
XX ligands to facilitate drug design and development. The present  
XX sequence is a complementary peptide provided in the specification.

XX Sequence 10 AA;

Query Match 76.5%; Score 13; DB 22; Length 10;  
Best Local Similarity 50.0%; Pred. No. 4.4e+02;  
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7  
Db 2 ESTALL 7

RESULT 3  
ABU08967  
ID ABU08967 standard; peptide; 11 AA.

XX AC ABU08967;  
XX 10-JUN-2003 (first entry)  
XX Human complement C3f derived peptide.

XX Human; complement; C3f; SELDI; myocardial infarction;  
XX surface enhanced laser desorption ionisation; intracerebral haemorrhage;  
XX congestive heart failure; mass spectroscopy; immunoassay;  
XX radioimmunoassay; enzyme-linked immunosorbent assay; ELISA;  
XX fluorescent immunoassay.

XX Homo sapiens.

XX US2002161184-A1.

XX 31-OCT-2002.

XX 30-APR-2001; 2001US-0845715.

XX 30-APR-2001; 2001US-0845715.

XX (JACK/) JACKOWSKI G.

XX (THAT/) THATCHER B.

XX (MARS/) MARSHALL J.

XX (YANT/) YANTHA J.

XX (VREE/) VREES T.

XX Jackowski G, Thatcher B, Marshall J, Yantha J, Vrees T;

XX WPI; 2003-340874/32.

XX Biopolymer marker, useful in diagnosing disease states including  
PT myocardial infarction and intracerebral hemorrhage, comprises  
PT complement C3f fragment with a specified molecular weight -

XX Claim 1; Page 7; 11pp; English.

XX The invention relates to a biopolymer marker comprises a complement C3f  
XX fragment appearing as ABU08967 having a molecular weight of 1348 daltons.  
XX The fragment was isolated from samples from affected individuals using  
XX SELDI (surface enhanced laser desorption ionisation) and a protein chip.

CC The marker is useful in methods for diagnosing a disease state such as  
 CC myocardial infarction, intracerebral hemorrhage or congestive heart  
 CC failure. The methods used include mass spectroscopy or immunoassays, e.g.  
 CC radioimmunoassay, enzyme-linked immunosorbent assay (ELISA) or  
 CC fluorescent immunoassays. The invention enables a diagnostician to  
 CC characterise the presence or absence of the disease state relative to  
 CC recognition of the presence or absence of the biopolymer, respectively.  
 CC The present sequence is the human complement C3f derived peptide of  
 CC the invention.  
 XX  
 XX SQ Sequence 11 AA;

Query Match 76.5%; Score 13; DB 24; Length 11;  
 Best Local Similarity 50.0%; Pred. NO. 4.8e+02;  
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7  
 |  
 |  
 |  
 6 ESASLL 11

RESULT 4  
 ABG72971  
 ID ABG72971 standard; Peptide; 12 AA.

AC ABG72971;  
 XX  
 XX 27-FEB-2003 (first entry)  
 XX DT  
 XX Human complement C3f fragment.  
 XX DE  
 XX Human; biopolymer marker; myocardial infarction; C3f; syndrome X;  
 KW congestive heart failure; intracerebral haemorrhage; complement.  
 KW  
 XX Homo sapiens.

XX Key Location/Qualifiers  
 FH Misc-difference 1 /note= "His is optionally absent"  
 FT Misc-difference 12  
 FT Misc-difference 12 /note= "Arg is optionally absent"

XX WO200288731-A2.  
 XX PD 07-NOV-2002.  
 XX PP 29-APR-2002; 2002WO-CA00632.

XX 30-APR-2001; 2001US-0845731.  
 XX PA (SYNX-) SYN.X PHARMA INC.  
 XX PI Jackowski G, Thatcher B, Marshall J, Yantha J, Vrees T;  
 XX WPI; 2003-111908/10.

XX New biopolymer markers useful for indicating one particular disease  
 PT state such as myocardial infarction, congestive heart failure and  
 PT intracerebral hemorrhage -  
 XX

PS Claim 1; Page 30; 30pp; English.

XX The invention relates to a biopolymer marker having a sequence  
 CC useful in indicating at least one particular disease state. The  
 CC biopolymer marker is useful for indicating one particular disease state  
 CC such as syndrome X, myocardial infarction, congestive heart failure and  
 CC intracerebral haemorrhage. The marker is released into the circulation  
 CC where it may be present in the blood or any blood product. The presence  
 CC of each marker is determined using antibodies specific for each of the  
 CC markers and detecting specific binding of each antibody to its  
 CC respective marker. The present sequence represents a biopolymer marker  
 CC which is characterised as a human C3f fragment from the complement  
 CC system.

XX SQ Sequence 12 AA;

Query Match 76.5%; Score 13; DB 24; Length 12;  
 Best Local Similarity 50.0%; Pred. NO. 5.2e+02;  
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7  
 |  
 |  
 |  
 6 ESASLL 11

RESULT 5  
 AAM00721  
 ID AAM00721 standard; Peptide; 14 AA.

XX AAM00721;  
 XX 01-OCT-2001 (first entry)  
 XX DT  
 XX Human protein fragment SEQ ID NO: 1271.  
 XX DE  
 XX Human; single nucleotide polymorphism; SNP; paternity test;  
 KW forensic test; aberrant protein expression.  
 KW  
 XX Homo sapiens.

XX WO200151670-A2.  
 XX PD 19-JUL-2001.  
 XX PP 05-JAN-2001; 2001WO-US00322;  
 XX PR 07-JAN-2000; 2000US-0174962.

XX (CURA-) CURAGEN CORP.  
 XX Shimkets RA, Leach MD;  
 XX WPI; 2001-451871/48.  
 XX DR N-PSDB; AAN89840.

XX Isolated human polynucleotides containing single nucleotide  
 PT polymorphisms; useful for the treatment and diagnosis of e.g. cancer,  
 PT infection and diabetes -  
 XX

PS Disclosure; Page 467; 475pp; English.

XX The present invention relates to human nucleic acids containing single  
 CC nucleotide polymorphisms (SNPs). These can be used in forensic and  
 CC paternity tests, and to aid in the treatment of diseases associated with  
 CC aberrant protein expression, including cancer, amyloidosis, diabetes,  
 CC Alzheimer's disease, Down's syndrome, oedema, lupus (SLE), vasculitis,  
 CC glomerulonephritis, haemolytic anaemia, thrombocytopaenia, arthritis,  
 CC meningitis, muscular disorders, dementia, neurological diseases, tubercu-  
 CC sclerosis, male infertility, hypercalcaemia, blood pressure disorders,  
 CC osteoporosis, pathogenic infections, hypercholesterolaemia, obesity or  
 CC autoimmunity. The present sequence is a peptide encoded by a  
 CC polymorphism-containing oligonucleotide fragment of the invention.

XX SQ Sequence 14 AA;

Query Match 76.5%; Score 13; DB 22; Length 14;  
 Best Local Similarity 50.0%; Pred. NO. 6e+02;  
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7  
 |  
 |  
 |  
 1 ESSTLL 6

RESULT 6  
 AAM00722

ID XX AAM00722 standard; Peptide; 14 AA.  
 AC AAM00722;  
 DT 01-OCT-2001 (first entry)  
 DE Human protein fragment SEQ ID NO: 1272.  
 KW Human; single nucleotide polymorphism; SNP; paternity test;  
 KW forensic test; aberrant protein expression.  
 XX Homo sapiens.  
 OS WO200151670-A2.  
 PN 19-JUL-2001.  
 PD 05-JAN-2001; 2001WO-US00322.  
 PF 07-JAN-2000; 2000US-0174962.  
 PR (CURA-) CURAGEN CORP.  
 XX Shinketsu RA, Leach MD;  
 PI WPI: 2001-451871/48.  
 DR N-PSDB; AAH89841.  
 XX Isolated human polynucleotides containing single nucleotide  
 PT polymorphisms, useful for the treatment and diagnosis of e.g. cancer,  
 PT infection and diabetes.  
 XX Disclosure; Page 467; 475pp; English.  
 XX The present invention relates to human nucleic acids containing single  
 CC nucleotide polymorphisms (SNPs). These can be used in forensic and  
 CC paternity tests, and to aid in the treatment of diseases associated with  
 CC aberrant protein expression, including cancer, amyloidosis, diabetes,  
 CC Alzheimer's disease, Down's syndrome, oedema, lupus (SLE), vasculitis,  
 CC glomerulonephritis, haemolytic anaemia, thrombocytopenia, arthritis,  
 CC meningitis, muscular disorders, dementia, neurological diseases, tuberculous  
 CC sclerosis, male infertility, hypercalcaemia, blood pressure disorders,  
 CC osteoporosis, pathogenic infections, hypercholesterolaemia, obesity or  
 CC autoimmunity. The present sequence is a peptide encoded by a  
 CC polymorphism-containing oligonucleotide fragment of the invention.  
 XX SQ Sequence 14 AA;  
 Query Match 76.5%; Score 13; DB 22; Length 14;  
 Best Local Similarity 50.0%; Pred. No. 6e+02; 3; Indels 0; Gaps 0;  
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 QY 2 EXXXLL 7  
 Db 1 ESSTLL 6  
 RESULT 7  
 ABP58087  
 ID ABP58087 standard; Peptide; 14 AA.  
 AC ABP58087;  
 XX 07-MAR-2003 (first entry)  
 DE Complement C3f fragment, specific disease marker.  
 KW Complement C3f; human; marker; myocardial infarction;  
 KW intracerebral haemorrhage; congestive heart failure; cardiac;  
 KW cerebroprotective; haemostatic.  
 XX Homo sapiens.  
 OS

PN WO200288727-A2.  
 XX 07-NOV-2002.  
 PD 29-APR-2002; 2002WO-CA00617.  
 PF 30-APR-2001; 2001US-0845730.  
 PR (SYNX-) SYN.X PHARMA INC.  
 PA Jackowski G, Thatcher B, Vrees T, Yantha J, Marshall J;  
 XX WPI: 2003-111907/10.  
 DR Biopolymer marker useful for evidencing, categorizing or regulating at  
 PT least one disease state, e.g. congestive heart failure, myocardial  
 PT infarction or intracerebral hemorrhage.  
 XX Claim 1; Page 30; 30pp; English.  
 CC The present sequence is that of a biopolymer marker that is useful  
 CC in indicating at least one particular disease state. The marker  
 CC is a C3f fragment from the complement system having a mol.wt. of  
 CC about 1690 Da. The marker is indicative of an individual suffering  
 CC from myocardial infarction, intracerebral haemorrhage or congestive  
 CC heart failure. The invention involves the use of a combination of  
 CC preparatory steps in conjunction with mass spectroscopy and  
 CC time-of-flight detection procedures to maximise the diversity of  
 CC biopolymers which are verifiable within a particular sample. The  
 CC cohort of biopolymers verified within such a sample is then viewed  
 CC with reference to their ability to evidence at least one particular  
 CC disease state, thereby enabling a diagnostician to gain the ability  
 CC to characterise either the presence or absence of a disease state  
 CC relative to recognition of the presence and/or absence of the  
 CC biopolymer. Diagnostic assay kits including the present biopolymer  
 CC marker are claimed, and a process for regulating a disease state  
 CC by controlling the presence or absence of the biopolymer marker is  
 CC also claimed.  
 XX SQ Sequence 14 AA;  
 Query Match 76.5%; Score 13; DB 24; Length 14;  
 Best Local Similarity 50.0%; Pred. No. 6e+02; 3; Indels 0; Gaps 0;  
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 QY 2 EXXXLL 7  
 Db 9 ESASLL 14  
 RESULT 8  
 ABG75803  
 ID ABG75803 standard; Peptide; 15 AA.  
 XX AC ABG75803;  
 XX 07-MAY-2003 (first entry)  
 DE Congestive heart failure C3f fragment marker biopolymer.  
 KW Biopolymer; disease state; congestive heart failure; asymptomatic;  
 KW diabetes; kidney failure; heart failure; Syndrome X; heart attack;  
 KW stroke; C3f; complement system.  
 XX Unidentified.  
 OS US2002160419-A1.  
 PN 31-OCT-2002.  
 PD 30-APR-2001; 2001US-0845739.  
 PF 30-APR-2001; 2001US-0845739.  
 PR

XX (JACK/) JACKOWSKI G.  
 PA (THAT/) THATCHER B.  
 PA (MARS/) MARSHALL J.  
 PA (YANT/) YANTHA J.  
 PA (VRES/) VREES T.  
 XX Jackowski G, Thatcher B, Marshall J, Yantha J, Vrees T;  
 XX WPI; 2003-246643/25.  
 XX A novel biopolymer marker useful in indicating at least one particular  
 PT disease state particularly congestive heart failure -  
 XX  
 XX Claim 1; Page 7; 10pp; English.  
 XX The invention discloses a biopolymer marker which is useful in indicating  
 CC at least one particular disease state, particularly congestive heart  
 CC failure. The polymer marker is also useful for characterizing disease  
 CC states. The polymer marker allows physicians to identify asymptomatic  
 CC patients before they suffer an irreversible event such as diabetes,  
 CC kidney failure and heart failure, and enable effective disease management  
 CC and preventative medicine. Additionally, the specific diagnostic tests  
 CC which evolve using the biopolymer provide a tool for rapidly and  
 CC accurately diagnosing acute Syndrome X such as heart attack and stroke,  
 CC and facilitate treatment. The sequence presented is the congestive heart  
 CC failure specific C3f fragment from the complement system, which has a  
 CC molecular weight of about 1793 Daltons, marker biopolymer.  
 XX  
 XX Sequence 15 AA;

Query Match 76.5%; Score 13; DB 24; Length 15;  
 Best Local Similarity 50.0%; Pred. No. 6.4e+02;  
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 EXXXLL 7  
 Db 10 ESASLL 15

RESULT 9  
 AAE05786  
 ID AAE05786 standard; peptide; 17 AA.  
 AC AAE05786;  
 XX  
 XX 24-SEP-2001 (first entry)  
 XX Human calcium binding protein 86 (CBP86) progesterone receptor motif #2.  
 XX Human; sperm specific protein; calcium binding protein 86; CBP86;  
 KW tyrosine phosphorylation; capacitation; male contraceptive agent;  
 KW progesterone receptor motif.  
 XX Homo sapiens.  
 OS WO200153338-A2.  
 PN  
 XX  
 XX 26-JUL-2001.  
 PD  
 XX  
 XX 19-JAN-2001; 2001WO-US01715.  
 PF  
 XX  
 XX 19-JAN-2000; 2000US-0176887.  
 PR  
 XX (UYVI-) UNIV VIRGINIA PATENT FOUND.  
 PA  
 XX Herr JC, Buer S, Mandal A, Wolkowicz M, Naaby-Hansen S;  
 PI WPI; 2001-465366/50.  
 XX  
 XX New calcium binding protein 86 (CBP86) polypeptides and polynucleotides  
 PT useful for isolating ligands that bind to the CBP86 polypeptides, or as  
 PT marker for determining the extent of capacitation of sperm cells

PT present in a sperm sample -  
 XX  
 PS Disclosure; Page 9; 69pp; English.  
 XX

CC The invention relates to human sperm specific protein designated  
 CC calcium binding protein 86 (CBP86). The protein exhibits  
 CC increased tyrosine phosphorylation and also increased calcium  
 CC binding isoforms during in vitro capacitation. CBP86 is expressed post-  
 CC meiotically and localised in the sperm flagellum. The CBP86 polypeptide  
 CC may be used to isolate ligands that bind to it under physiological  
 CC conditions, and as marker for determining the extent of capacitation of  
 CC sperm cells. CBP86 antagonists may be used to interfere with the  
 CC capacitation of vertebrate sperm, and thus used as contraceptive agents.  
 CC Antibodies against the CBP86 protein can be used for the diagnosis of  
 CC CBP86, or in assays to monitor patients being treated with CBP86  
 CC agonists, antagonists or inhibitors. The present sequence is a  
 CC progesterone receptor motif found in human CBP86.

XX Sequence 17 AA;

Query Match 76.5%; Score 13; DB 22; Length 17;  
 Best Local Similarity 50.0%; Pred. No. 7.1e+02;  
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 EXXXLL 7  
 Db 2 EATALL 7

RESULT 10  
 ABU08617  
 ID ABU08617 standard; peptide; 17 AA.  
 XX  
 AC ABU08617;  
 XX  
 XX 23-MAY-2003 (first entry)  
 DT  
 XX  
 DE Disease specific biopolymer marker #1.  
 XX  
 KW Biopolymer marker; type II diabetes; immunoassay.  
 XX Homo sapiens.

OS  
 XX US2002160532-A1.  
 PN  
 XX 31-OCT-2002.  
 PD

PF 30-APR-2001; 2001US-0846346.  
 XX  
 XX 30-APR-2001; 2001US-0846346.  
 PR

XX (JACK/) JACKOWSKI G.  
 PA (THAT/) THATCHER B.  
 PA (MARS/) MARSHALL J.  
 PA (YANT/) YANTHA J.  
 PA (VRES/) VREES T.

PI Jackowski G, Thatcher B, Marshall J, Yantha J, Vrees T;  
 WPI; 2003-328370/31.

XX Biopolymer marker useful in indicating disease state, in particular  
 PT type II diabetes and as antigens in immunoassays for detecting  
 PT individuals suffering from disease known to be evidenced by marker  
 PT sequence -

XX Claim 1; Page 7; 10pp; English.

PS The invention describes a biopolymer marker (I) useful in indicating at  
 CC least one particular disease state. (I) is useful for indicating a  
 CC disease state, in particular type II diabetes. The marker sequences are  
 CC useful as antigens in immunoassays for the detection of those individuals

CC suffering from the disease known to be evidenced by the marker sequence.  
 CC (1) provides an efficient diagnostic tool for rapidly and accurately  
 CC diagnosing disease states such as type II diabetes. This is the  
 CC amino acid sequence of a biopolymer marker.

SQ Sequence 17 AA;

Query Match 76.5%; Score 13; DB 24; Length 17;  
 Best Local Similarity 50.0%; Pred. No. 7.1e+02;  
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 EXXXLL 7  
 DB 11 ESASLL 16

RESULT 11

ABB82769  
 ID ABB82769 standard; peptide; 17 AA.

ABB82769;

DT 18-MAR-2003 (first entry)

DE Congestive heart failure indicative biopolymer marker.

XX Biopolymer; marker; C3f; complement system; congestive heart failure;  
 human.

OS Homo sapiens.

XX WO200288717-A2.

XX 07-NOV-2002.

XX 25-APR-2002; 2002WO-CA00578.

XX 30-APR-2001; 2001US-0845736.

XX (SYNX-) SYN.X PHARMA INC.

PI Jackowski G, Thatcher B, Marshall J, Yantha J, Vrees T;

XX WPI; 2003-120486/11.

PT Use of biopolymer marker for evidencing, categorizing or regulating at  
 PT least one disease state, e.g. congestive heart failure

XX Claim 1; Fig 1; 27pp; English.

CC The present sequence represents a biopolymer marker of the invention and  
 CC is a disease specific marker. The marker is characterised as a C3f  
 CC fragment from the complement system having a molecular weight of about  
 CC 2056 daltons. The biopolymer marker identified is useful for evidencing,  
 CC categorizing or regulating at least one disease state, preferably  
 CC congestive heart failure.

SQ Sequence 17 AA;

Query Match 76.5%; Score 13; DB 24; Length 17;  
 Best Local Similarity 50.0%; Pred. No. 7.1e+02;  
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 EXXXLL 7  
 DB 11 ESASLL 16

RESULT 12

ABU08618  
 ID ABU08618 standard; peptide; 18 AA.

XX ABU08618;

XX 23-MAY-2003 (first entry)  
 DT Disease specific biopolymer marker #2.  
 DE Biopolymer marker; type II diabetes; immunoassay.  
 XX Homo sapiens.  
 XX US2002160532-A1.  
 XX 31-OCT-2002.  
 XX 30-APR-2001; 2001US-0846346.  
 XX 30-APR-2001; 2001US-0846346.  
 XX (JACK/) JACKOWSKI G.  
 XX (THAT/) THATCHER B.  
 XX (MARS/) MARSHALL J.  
 XX (YANT/) YANTHA J.  
 XX (VREE/) VREES T.  
 XX Jackowski G, Thatcher B, Marshall J, Yantha J, Vrees T;  
 XX WPI; 2003-328370/31.  
 XX Biopolymer marker useful in indicating disease state, in particular  
 XX type II diabetes and as antigens in immunoassays for detecting  
 XX individuals suffering from disease known to be evidenced by marker  
 XX sequence  
 XX Disclosure; Fig 2; 10pp; English.  
 XX The invention describes a biopolymer marker (I) useful in indicating at  
 XX least one particular disease state. (I) is useful for indicating a  
 XX disease state, in particular type II diabetes. The marker sequences are  
 XX useful as antigens in immunoassays for the detection of those individuals  
 XX suffering from the disease known to be evidenced by the marker sequence.  
 XX (I) provides an efficient diagnostic tool for rapidly and accurately  
 XX diagnosing disease states such as type II diabetes. This is the  
 XX amino acid sequence of a biopolymer marker.  
 XX Sequence 18 AA;  
 SQ Query Match 76.5%; Score 13; DB 24; Length 18;  
 Best Local Similarity 50.0%; Pred. No. 7.5e+02;  
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 OY 2 EXXXLL 7  
 DB 12 ESASLL 17  
 RESULT 13  
 AAR97974  
 ID AAR97974 standard; peptide; 19 AA.  
 XX AAR97974;  
 AC AAR97974;  
 XX 12-NOV-1996 (first entry)  
 DT Antigenic fragment (p49) used in fusion protein construct.  
 XX Fusion protein; carrier segment; hydrophobic; insoluble; antibody;  
 XX immunopurification; antigenicity; ligand.  
 XX Synthetic.  
 XX WO9615249-A1.  
 XX 23-MAY-1996.  
 XX

PF 13-NOV-1995; 95WO-US14518.

PR 10-NOV-1994; 94US-0338382.

PA (PROM-) PROMEGA CORP.

PI Haak-frendscho M, Knuth MW, Lesley SA, Shultz JW;

PI Villars CE;

XX WPI; 1996-259849/26.

XX Hydrophobic, sparingly soluble fusion protein carrier segment - used  
PT in immunopurification of antibodies and prodn. of immunologically  
PT active peptide(s) and vaccines

PS Example 2; Page 63; 121pp; English.

XX Fusion protein carrier segments comprising non-naturally occurring  
CC hydrophobic sparingly soluble amino acid sequences may be used in a  
CC process for immunopurifying antibodies from a pool of antibody  
CC material. The carrier segments have reduced antigenicity but when  
CC injected into animals will adequately display an attached ligand.  
CC The carrier segment also lacks many cleavage sites recognised by  
CC host proteases and is soluble in some solutions, allowing cleavage  
CC by chemicals and enzymes.

XX Sequence 19 AA;

Query Match 76.5%; Score 13; DB 17; Length 19;

Best Local Similarity 50.0%; Pred. No. 7.9e+02;

Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 EXXLL 7

DB 1 RTSSL 6

RESULT 14

AAE05792

ID AAE05792 standard; peptide; 20 AA.

AC AAE05792;

XX 24-SEP-2001 (first entry)

XX Human calcium binding protein 86 (CBP86) peptide #4.

XX Human; sperm specific protein; calcium binding protein 86; CBP86;  
XX tyrosine phosphorylation; capacitation; male contraceptive agent.

XX Homo sapiens.

OS WO2000153338-A2.

XX 26-JUL-2001.

XX 19-JAN-2001; 2001WO-US01715.

XX 19-JAN-2000; 2000US-0176887.

XX (UTVI-) UNIV VIRGINIA PATENT FOUND.

XX Herr JC, Buer S, Mandal A, Wolkowicz M, Naaby-Hansen S;

XX WPI; 2001-465366/50.

XX New calcium binding protein 86 (CBP86) polypeptides and polynucleotides  
XX useful for isolating ligands that bind to the CBP86 polypeptides, or as  
XX marker for determining the extent of capacitation of sperm cells  
XX present in a sperm sample -

XX Example 1; Page 16; 69pp; English.

CC The invention relates to human sperm specific protein designated  
CC calcium binding protein 86 (CBP86). The protein exhibits  
CC increased tyrosine phosphorylation and also increased calcium  
CC binding isoforms during in vitro capacitation. CBP86 is expressed post-  
CC meiotically and localised in the sperm flagellum. The CBP86 polypeptide  
CC may be used to isolate ligands that bind to it under physiological  
CC conditions, and as marker for determining the extent of capacitation of  
CC sperm cells. CBP86 antagonists may be used to interfere with the  
CC capacitation of vertebrate sperm, and thus used as contraceptive agents.  
CC Antibodies against the CBP86 protein can be used for the diagnosis of  
CC conditions or diseases characterised by expression or overexpression of  
CC CBP86, or in assays to monitor patients being treated with CBP86  
CC agonists, antagonists or inhibitors. The present sequence is  
CC human CBP86 peptide.

XX Sequence 20 AA;

Query Match 76.5%; Score 13; DB 22; Length 20;

Best Local Similarity 50.0%; Pred. No. 8.2e+02;

Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 EXXLL 7

DB 10 EATALL 15

RESULT 15

AAE033790

ID AAE033790 standard; Peptide; 25 AA.

XX AAE033790;

XX 02-FEB-2001 (first entry)

XX Human secreted protein BLAST search protein SEQ ID NO: 134.

XX Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;  
XX anti-allergic; hepatotropic; antidiabetic; anti-inflammatory; antitumor;  
XX vulnary; anticonvulsant; antibacterial; antifungal; antiparasitic;  
XX cardiac; gene therapy; cancer; immune disorder; cardiovascular disorder;  
XX neurological disease; infection; human; secreted protein.

XX Homo sapiens.

XX WO2000056753-A1.

XX 28-SEP-2000.

XX 16-MAR-2000; 2000WO-US06785.

XX 23-MAR-1999; 99US-0126051.

XX 10-DEC-1999; 99US-0169906.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Ruben SM, Komatsoulis G;

XX WPI; 2000-594570/56.

XX Nucleic acid molecules encoding human secreted proteins, used in  
XX preventing, treating or ameliorating a disorder -

XX Disclosure; Page 405; 410pp; English.

XX The invention relates to the isolation of genes AAC59277-C59325 encoding  
XX the human secreted proteins AAE033790-833764. The sequence is used as a  
XX query sequence for doing BLASTX searches to identify homologous  
XX sequences. The genes and proteins are useful for preventing,  
XX ameliorating or treating medical conditions, e.g. by protein or gene  
XX therapy. The genes are isolated from a range of human tissues disclosed  
XX in the specification. The nucleic acids, proteins, antibodies and  
XX (ant)agonists are useful in the diagnosis, treatment and prevention of:  
XX (a) cancer, e.g. breast and ovarian cancer, and other cancers of the

CC adrenal gland, bone, bone marrow, breast, gastrointestinal  
 CC tract, liver, lung, or urogenital; (b) immune disorders e.g. Addison's  
 CC disease, allergies, autoimmune haemolytic anaemia, autoimmune  
 CC thyroiditis, diabetes mellitus, Crohn's disease, multiple sclerosis,  
 CC rheumatoid arthritis and ulcerative colitis; (c) cardiovascular  
 CC disorders such as myocardial ischaemias; (d) wound healing; (e)  
 CC neurological diseases e.g. cerebral anoxia and epilepsy; and (f)  
 CC infectious diseases such as viral, bacterial, fungal and parasitic  
 CC infections.

XX Sequence 25 AA;

Query Match 76.5%; Score 13; DB 21; Length 25;

Best Local Similarity 50.0%; Pred. No. 16+03;

Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7

Db 11 ETSSLL 16

LT 16

AM88391

ID AM88391 standard; Protein; 26 AA.

XX AC AM88391;

XX DT 07-NOV-2001 (first entry)

XX Human immune/haematopoietic antigen SEQ ID NO:15984.

XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;  
 KW cytostatic; gene therapy; vaccine; metastasis.

XX OS Homo sapiens.

XX WO200157182-A2.

XX PD 09-AUG-2001.

XX PF 17-JAN-2001; 2001WO-US01354.

XX 31-JAN-2000; 2000US-0179065.

XX 04-FEB-2000; 2000US-0180628.

XX 24-FEB-2000; 2000US-0184664.

XX 02-MAR-2000; 2000US-0186350.

XX 16-MAR-2000; 2000US-0189874.

XX 17-MAR-2000; 2000US-0190076.

XX 18-APR-2000; 2000US-0198123.

XX 19-MAY-2000; 2000US-0205515.

XX 07-JUN-2000; 2000US-0209467.

XX 28-JUN-2000; 2000US-0214886.

XX 30-JUN-2000; 2000US-0215135.

XX 07-JUL-2000; 2000US-0216647.

XX 07-JUL-2000; 2000US-0216890.

XX 11-JUL-2000; 2000US-0217487.

XX 11-JUL-2000; 2000US-0217496.

XX 14-JUL-2000; 2000US-0218290.

XX 26-JUL-2000; 2000US-0220963.

XX 26-JUL-2000; 2000US-0220964.

XX 14-AUG-2000; 2000US-0224518.

XX 14-AUG-2000; 2000US-0224519.

XX 14-AUG-2000; 2000US-0225213.

XX 14-AUG-2000; 2000US-0225214.

XX 14-AUG-2000; 2000US-0225266.

XX 14-AUG-2000; 2000US-0225267.

XX 14-AUG-2000; 2000US-0225268.

XX 14-AUG-2000; 2000US-0225270.

XX 14-AUG-2000; 2000US-0225447.

XX 14-AUG-2000; 2000US-0225757.

XX 14-AUG-2000; 2000US-0225758.

XX 14-AUG-2000; 2000US-0225759.

XX 18-AUG-2000; 2000US-0226279.

PR 22-AUG-2000; 2000US-0226681.  
 PR 22-AUG-2000; 2000US-0226688.  
 PR 22-AUG-2000; 2000US-0227182.  
 PR 23-AUG-2000; 2000US-0227009.  
 PR 30-AUG-2000; 2000US-0228924.  
 PR 01-SEP-2000; 2000US-0229287.  
 PR 01-SEP-2000; 2000US-0229343.  
 PR 01-SEP-2000; 2000US-0229344.  
 PR 01-SEP-2000; 2000US-0229345.  
 PR 05-SEP-2000; 2000US-0229509.  
 PR 06-SEP-2000; 2000US-0229513.  
 PR 06-SEP-2000; 2000US-0230437.  
 PR 08-SEP-2000; 2000US-0230438.  
 PR 08-SEP-2000; 2000US-0231242.  
 PR 08-SEP-2000; 2000US-0231243.  
 PR 08-SEP-2000; 2000US-0231244.  
 PR 08-SEP-2000; 2000US-0231413.  
 PR 08-SEP-2000; 2000US-0231414.  
 PR 08-SEP-2000; 2000US-0232080.  
 PR 12-SEP-2000; 2000US-0232081.  
 PR 14-SEP-2000; 2000US-0231968.  
 PR 14-SEP-2000; 2000US-0232397.  
 PR 14-SEP-2000; 2000US-0232398.  
 PR 14-SEP-2000; 2000US-0232399.  
 PR 14-SEP-2000; 2000US-0232400.  
 PR 14-SEP-2000; 2000US-0232401.  
 PR 14-SEP-2000; 2000US-0233063.  
 PR 14-SEP-2000; 2000US-0233064.  
 PR 14-SEP-2000; 2000US-0233065.  
 PR 21-SEP-2000; 2000US-0234223.  
 PR 21-SEP-2000; 2000US-0234274.  
 PR 25-SEP-2000; 2000US-0234997.  
 PR 25-SEP-2000; 2000US-0234998.  
 PR 26-SEP-2000; 2000US-0235484.  
 PR 27-SEP-2000; 2000US-0235834.  
 PR 27-SEP-2000; 2000US-0235836.  
 PR 29-SEP-2000; 2000US-0236327.  
 PR 29-SEP-2000; 2000US-0236367.  
 PR 29-SEP-2000; 2000US-0236368.  
 PR 29-SEP-2000; 2000US-0236369.  
 PR 29-SEP-2000; 2000US-0236370.  
 PR 02-OCT-2000; 2000US-0236802.  
 PR 02-OCT-2000; 2000US-0237037.  
 PR 02-OCT-2000; 2000US-0237038.  
 PR 02-OCT-2000; 2000US-0237039.  
 PR 13-OCT-2000; 2000US-0239935.  
 PR 13-OCT-2000; 2000US-0239937.  
 PR 20-OCT-2000; 2000US-0240960.  
 PR 20-OCT-2000; 2000US-0241221.  
 PR 20-OCT-2000; 2000US-0241785.  
 PR 20-OCT-2000; 2000US-0241786.  
 PR 20-OCT-2000; 2000US-0241787.  
 PR 20-OCT-2000; 2000US-0241808.  
 PR 20-OCT-2000; 2000US-0241809.  
 PR 20-OCT-2000; 2000US-0241826.  
 PR 01-NOV-2000; 2000US-0244617.  
 PR 08-NOV-2000; 2000US-0246474.  
 PR 08-NOV-2000; 2000US-0246475.  
 PR 08-NOV-2000; 2000US-0246476.  
 PR 08-NOV-2000; 2000US-0246477.  
 PR 08-NOV-2000; 2000US-0246478.  
 PR 08-NOV-2000; 2000US-0246523.  
 PR 08-NOV-2000; 2000US-0246524.  
 PR 08-NOV-2000; 2000US-0246525.  
 PR 08-NOV-2000; 2000US-0246526.  
 PR 08-NOV-2000; 2000US-0246527.  
 PR 08-NOV-2000; 2000US-0246528.  
 PR 08-NOV-2000; 2000US-0246532.  
 PR 08-NOV-2000; 2000US-0246609.  
 PR 08-NOV-2000; 2000US-0246610.  
 PR 08-NOV-2000; 2000US-0246611.  
 PR 08-NOV-2000; 2000US-0246613.





ID AAO11963 standard; Protein; 34 AA.  
 AC AAO11963;  
 XX  
 DT 06-NOV-2001 (first entry)  
 XX  
 DE Human polypeptide SEQ ID NO 25855.  
 XX  
 KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;  
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;  
 KW nervous system disorders; arthritis; inflammation.  
 OS Homo sapiens.  
 XX  
 PN WO200164835-A2.  
 XX  
 PD 07-SEP-2001.  
 XX  
 PR 26-FEB-2001; 2001WO-US04927.  
 XX  
 PR 28-FEB-2000; 2000US-0515126.  
 XX  
 PR 18-MAY-2000; 2000US-0577409.  
 XX  
 PA (HYSE-) HYSEQ INC.  
 XX  
 PI Tang YT, Liu C, Drmanac RT;  
 XX  
 DR WPI; 2001-514838/56.  
 DR N-PSDB; AAI191894.  
 XX  
 PT Isolated nucleic acids and polypeptides, useful for preventing  
 PT diagnosing and treating e.g. leukaemia, inflammation and immune  
 PT disorders -  
 XX  
 PS Claim 20; SEQ ID NO 25855; 1399pp + Sequence Listing; English.  
 XX  
 CC The invention relates to human polynucleotides (AAI19941-AAI193841) and  
 CC the encoded proteins (AAO00010-AAO13910) that exhibit activity relating to  
 CC cytokine, cell proliferation or cell differentiation or which may induce  
 CC production of other cytokines in other cell populations. The  
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
 CC peptide therapy. The polypeptides have various cytokine-like activities,  
 CC e.g. stem cell growth factor activity, haematopoiesis regulating  
 CC activity, tissue growth factor activity, immunomodulatory activity and  
 CC activin/inhibin activity and may be useful in the diagnosis and/or  
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and  
 CC inflammation.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 SQ Sequence 34 AA;  
 Query Match 76.5%; Score 13; DB 22; Length 34;  
 Best Local Similarity 50.0%; Pred. No. 1.3e+03;  
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 QY 2 EXXXLL 7  
 DB 16 ETTSL 21  
 RESULT 19  
 ID AAR28125 standard; Protein; 42 AA.  
 XX  
 AC AAR28125;  
 XX  
 DT 25-MAR-2003 (updated)  
 DT 12-MAR-1993 (first entry)  
 XX  
 DE Peptide sequence encoded by GUS:NPT-II fusion junction.

XX Genetic selection; genetic marker; gene recovery.  
 KW Synthetic.  
 OS  
 PN WO9217593-A1.  
 XX  
 PD 15-OCT-1992.  
 XX  
 PR 30-MAR-1992; 92WO-CA00139.  
 XX  
 PR 28-MAR-1991; 91US-0676432.  
 XX  
 PA (CANA) NAT RES COUNCIL CANADA.  
 XX  
 PI Crosby WL, Datla RSS, Hammerlindl JK, Selvaraj G;  
 XX  
 DR WPI; 1992-366262/44.  
 DR N-PSDB; AAQ29916.  
 XX  
 PT Fused gene with characteristics of component parts - comprising a  
 PT gene conferring conditional growth advantage and a marker gene,  
 PT used as a probe to select and isolate genetic elements  
 XX  
 PS Disclosure; Page 30; 53pp; English.  
 XX  
 CC The sequence is that encoded by the region surrounding the junction  
 CC of GUS (beta-glucuronidase) and NPT-II (neomycin phosphotransferase-II)  
 CC in pGK314. See also AAR28127.  
 CC (Updated on 25-MAR-2003 to correct PN field.)  
 XX  
 SQ Sequence 42 AA;  
 Query Match 76.5%; Score 13; DB 13; Length 42;  
 Best Local Similarity 50.0%; Pred. No. 1.6e+03;  
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 QY 2 EXXXLL 7  
 DB 14 ESTTL 19  
 RESULT 20  
 ID AAO10031 standard; Protein; 46 AA.  
 XX  
 AC AAO10031;  
 XX  
 DT 06-NOV-2001 (first entry)  
 XX  
 DE Human polypeptide SEQ ID NO 23923.  
 XX  
 KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;  
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;  
 KW nervous system disorders; arthritis; inflammation.  
 OS Homo sapiens.  
 XX  
 PN WO200164835-A2.  
 XX  
 PD 07-SEP-2001.  
 XX  
 PR 26-FEB-2001; 2001WO-US04927.  
 XX  
 PR 28-FEB-2000; 2000US-0515126.  
 XX  
 PR 18-MAY-2000; 2000US-0577409.  
 XX  
 PA (HYSE-) HYSEQ INC.  
 XX  
 PI Tang YT, Liu C, Drmanac RT;  
 XX  
 DR WPI; 2001-514838/56.

DR N-PSDB; AA189962.

XX Isolated nucleic acids and polypeptides, useful for preventing  
 PT diagnosing and treating e.g. leukaemia, inflammation and immune  
 PT disorders -

XX Claim 20; SEQ ID NO 23923; 1399pp + Sequence Listing; English.

XX The invention relates to human polynucleotides (AA179941-AA193841) and  
 CC the encoded proteins (AA000010-AA013910) that exhibit activity elating to  
 CC cytokine, cell proliferation or cell differentiation or which may induce  
 CC production of other cytokines in other cell populations. The  
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
 CC peptide therapy. The polypeptides have various cytokine-like activities,  
 CC e.g. stem cell growth factor activity, haematopoiesis regulating  
 CC activity, tissue growth factor activity, immunomodulatory activity and  
 CC activin/inhibin activity and may be useful in the diagnosis and/or  
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and  
 CC inflammation.

XX Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 46 AA;

Query Match 76.5%; Score 13; DB 22; Length 46;  
 Best Local Similarity 50.0%; Pred. No. 1.7e+03;  
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7

DB 16 ETSSLL 21

RESULT 21

ABG50261  
 ID ABG50261 standard; Peptide; 53 AA.

XX AC ABG50261;

XX 25-FEB-2003 (first entry)

XX Human liver peptide, SEQ ID No 28909.

XX Human; liver; cirrhosis; hyperlipoproteinaemia; hyperlipidaemia;  
 XX hypercholesterolaemia; coronary heart disease.

XX Homo sapiens.

XX WO200157273-A2.

XX 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US00664.

XX 04-FEB-2000; 2000US-0180312.

XX 26-MAY-2000; 2000US-0207455.

XX 30-JUN-2000; 2000US-0608408.

XX 03-AUG-2000; 2000US-0632366.

XX 21-SEP-2000; 2000US-0234687.

XX 27-SEP-2000; 2000US-0236359.

XX 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-488898/53.

XX Human genome-derived single exon nucleic acid probes useful for  
 PT analysing gene expression in human adult liver -

XX Claim 27; SEQ ID No 28909; 658pp; English.

XX

CC The invention relates to a single exon nucleic acid probe (SENP) (I) for  
 CC measuring human gene expression in a sample derived from human adult  
 CC liver, comprising one of 13109 defined nucleotide sequences given in the  
 CC specification (or complements/ fragments). The probe hybridises at high  
 CC stringency to a nucleic acid molecule expressed in the human adult  
 CC liver. (I) may be used for predicting, measuring and displaying gene  
 CC expression in samples derived from human adult liver. The genes  
 CC identified may be involved in genetic liver diseases such as cirrhosis,  
 CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which  
 CC is associated with coronary heart disease. ABG47348-ABG59930 represent  
 CC human liver single exon encoded peptides of the invention.  
 CC Note: The sequence information for this patent does not appear in the  
 CC printed specification but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 53 AA;

Query Match 76.5%; Score 13; DB 22; Length 53;  
 Best Local Similarity 50.0%; Pred. No. 2e+03;  
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7

DB 32 EATALL 37

RESULT 22

ABB30233

XX ID ABB30233 standard; Peptide; 53 AA.

XX AC ABB30233;

XX 01-FEB-2002 (first entry)

XX Peptide #2884 encoded by breast cell single exon nucleic acid probe.

XX Human; microarray; single exon probe; gene expression; breast;  
 XX disease; cancer.

XX Homo sapiens.

XX WO200157271-A2.

XX 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US00662.

XX 04-FEB-2000; 2000US-0180312.

XX 26-MAY-2000; 2000US-0207456.

XX 30-JUN-2000; 2000US-0608408.

XX 03-AUG-2000; 2000US-0632366.

XX 21-SEP-2000; 2000US-0234687.

XX 27-SEP-2000; 2000US-0236359.

XX 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-496933/54.

XX New spatially-addressable set of single exon nucleic acid probes,  
 XX useful for measuring gene expression in sample derived from human  
 XX breast, comprises number of single exon nucleic acid probes -

XX Claim 27; SEQ ID NO 13201; 327pp + sequence listing; English.

XX The invention relates to a spatially-addressable set of single exon  
 CC nucleic acid probes for measuring gene expression in a sample derived  
 CC from human breast and BT 474 cells. The method involves contacting  
 CC the probes with a collection of detectably labelled nucleic acids  
 CC derived from mRNA of human breast, and then measuring the label

CC bound to each probe of the microarray. The probes are useful for  
CC verifying the expression of regions of genomic DNA predicted to  
CC encode proteins. They are useful for gene discovery, and for  
CC determining predisposition and/or prognosing breast disease. Gene  
CC expression analysis is useful for assessing the toxicity of chemical  
CC agents on cells. The microarray of this invention presents a far greater  
CC diversity of probes for measuring gene expression, with far less bias  
CC than expressed sequence tag microarrays. The method is suitable for  
CC rapid production of functional information from genomic sequence. The  
CC present sequence is a peptide encoded by a single exon nucleic acid  
CC probe of the invention.  
CC Note: The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
XX  
SQ Sequence 53 AA;  
Query Match 76.5%; Score 13; DB 22; Length 53;  
Best Local Similarity 50.0%; Pred. No. 2e+03; 3; Indels 0; Gaps 0;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
2 EXXXLL 7  
32 EATALL 37  
Db  
RESULT 23  
ABB35396  
ID ABB35396 standard; Peptide; 53 AA.  
XX ABB35396;  
XX  
XX  
DT 04-FEB-2002 (first entry)  
DE Peptide #2902 encoded by human foetal liver single exon probe.  
KW Human; foetal liver; gene expression; single exon nucleic acid probe.  
XX  
XX Homo sapiens.  
XX  
XX WO200157277-A2.  
XX  
XX PD 09-AUG-2001.  
XX  
XX PF 30-JAN-2001; 2001WO-US0669.  
XX  
XX 04-FEB-2000; 2000US-0180312.  
XX 26-MAY-2000; 2000US-0207456.  
XX 30-JUN-2000; 2000US-0608408.  
XX 03-AUG-2000; 2000US-0632366.  
XX 21-SEP-2000; 2000US-0234687.  
XX 27-SEP-2000; 2000US-0236359.  
XX 04-OCT-2000; 2000GB-0024263.  
XX  
XX (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX  
XX DR WPI; 2001-483447/52.  
XX  
XX PT Human genome-derived single exon nucleic acid probes useful for  
XX analyzing gene expression in human fetal liver -  
XX  
XX FS Claim 27; SEQ ID NO 28031; 639pp + sequence listing; English.  
XX  
XX The invention relates to a single exon nucleic acid probe for  
XX measuring human gene expression in a sample derived from human foetal  
XX liver. The single exon nucleic acid probes may be used for predicting,  
XX measuring and displaying gene expression in samples derived from human  
XX foetal liver. The present sequence is a peptide encoded by a single exon  
XX nucleic acid probe of the invention.  
XX Note: The sequence data for this patent did not form part of the  
XX printed specification, but was obtained in electronic format directly

CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
XX SQ Sequence 53 AA;  
Query Match 76.5%; Score 13; DB 22; Length 53;  
Best Local Similarity 50.0%; Pred. No. 2e+03; 3; Indels 0; Gaps 0;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
2 EXXXLL 7  
32 EATALL 37  
Db  
RESULT 24  
ABB17937  
ID ABB17937 standard; Protein; 53 AA.  
XX ABB17937;  
XX  
XX DT 23-JAN-2002 (first entry)  
DE Human nervous system related polypeptide SEQ ID NO 6594.  
XX  
XX Human; nootropic; neuroprotective; cytostatic; dermatological; virucide;  
KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnary;  
KW antiparkinsonian; antisticking; antianaemic; antiarthritic; cancer;  
KW antirheumatic; hepatotropic; cerebroprotective; antinflammatory;  
KW antiallergic; antidiabetic; antitumor; anticonvulsant; antifungal;  
KW antiparasitic; cardiant; immune disorder; cardiovascular disorder;  
KW neurological disease; infection; nephrotropic; gene therapy; vaccine.  
XX  
XX Homo sapiens.  
XX  
XX PN WO200159063-A2.  
XX  
XX PD 16-AUG-2001.  
XX  
XX PF 17-JAN-2001; 2001WO-US01334.  
XX  
XX 31-JAN-2000; 2000US-0179065.  
XX 04-FEB-2000; 2000US-0180628.  
XX 24-FEB-2000; 2000US-0184664.  
XX 02-MAR-2000; 2000US-0186350.  
XX 16-MAR-2000; 2000US-0189874.  
XX 17-MAR-2000; 2000US-0190076.  
XX 18-APR-2000; 2000US-0198123.  
XX 19-MAY-2000; 2000US-0205515.  
XX 07-JUN-2000; 2000US-0209467.  
XX 28-JUN-2000; 2000US-0214886.  
XX 30-JUN-2000; 2000US-0215135.  
XX 07-JUL-2000; 2000US-0216847.  
XX 07-JUL-2000; 2000US-0216880.  
XX 11-JUL-2000; 2000US-0217487.  
XX 11-JUL-2000; 2000US-0217496.  
XX 14-JUL-2000; 2000US-0218290.  
XX 26-JUL-2000; 2000US-0220963.  
XX 26-JUL-2000; 2000US-0220964.  
XX 14-AUG-2000; 2000US-0224518.  
XX 14-AUG-2000; 2000US-0224519.  
XX 14-AUG-2000; 2000US-0225213.  
XX 14-AUG-2000; 2000US-0225214.  
XX 14-AUG-2000; 2000US-0225266.  
XX 14-AUG-2000; 2000US-0225267.  
XX 14-AUG-2000; 2000US-0225268.  
XX 14-AUG-2000; 2000US-0225270.  
XX 14-AUG-2000; 2000US-0225447.  
XX 14-AUG-2000; 2000US-0225757.  
XX 14-AUG-2000; 2000US-0225758.  
XX 14-AUG-2000; 2000US-0225759.  
XX 18-AUG-2000; 2000US-0226279.  
XX 22-AUG-2000; 2000US-0226681.  
XX 22-AUG-2000; 2000US-0226686.  
XX 22-AUG-2000; 2000US-0227182.

PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231444.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 20-OCT-2000; 2000US-0242221.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.

PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250391.  
PR 01-DEC-2000; 2000US-0251160.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 11-DEC-2000; 2000US-0251990.  
PR 05-JAN-2001; 2000US-0254097.  
PR 05-JAN-2001; 2000US-0259678.  
XX (HUMA-) HUMAN GENOME SCI INC.  
PA  
XX  
XX  
PI Rosen CA, Barash SC, Ruben SM;  
XX  
XX WPI; 2001-541565/60.  
DR N-PSDB; ABA14263.  
XX  
XX  
PT Nucleic acids encoding 3224 human nervous system antigen polypeptides,  
PT useful for preventing, diagnosing and/or treating nervous system  
PT cancers and metastases -  
XX  
XX Claim 11; SEQ ID NO 6594; 1701pp + Sequence listing; English.  
PS  
XX The invention relates to novel genes (ABA11004-ABA21534) and proteins  
CC (ABA14578-ABA18001) useful for preventing, treating or ameliorating  
CC medical conditions e.g. by protein or gene therapy. The genes are  
CC isolated from a range of human tissues disclosed in the specification.  
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful  
CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast  
CC and ovarian cancer and other cancers of the adrenal gland, bone, bone  
CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;  
CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune  
CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's  
CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative  
CC colitis; (c) cardiovascular disorders such as myocardial ischaemia;  
CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and  
CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal  
CC and parasitic infections.  
CC Note: The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
XX Sequence 53 AA;

Query Match 76.5%; Score 13; DB 22; Length 53;  
Best Local Similarity 50.0%; Pred. No. 2e+03; 3; Indels 0; Gaps 0;  
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
Qy 2 EXXXLL 7  
Db 20 EASTLL 25

DE Human brain expressed single exon probe encoded protein SEQ ID NO: 28328.  
 XX Human; brain expressed exon; gene expression analysis; probe;  
 KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;  
 KW epilepsy; cancer.  
 XX Homo sapiens.  
 XX WO200157275-A2.  
 PN 09-AUG-2001.  
 PD 30-JAN-2001; 2001WO-US00667.  
 PF 04-FEB-2000; 2000US-0180312.  
 PR 26-MAY-2000; 2000US-0207456.  
 PR 30-JUN-2000; 2000US-0608408.  
 PR 03-AUG-2000; 2000US-0632366.  
 PR 21-SEP-2000; 2000US-0234687.  
 PR 27-SEP-2000; 2000US-0236359.  
 PR 04-OCT-2000; 2000GB-0024263.  
 XX (MOLE-) MOLECULAR DYNAMICS INC.  
 PA Penn SG, Hanzel DK, Chen W, Rank DR;  
 PI WPI; 2001-483446/52.  
 DR Single exon nucleic acid probes for analyzing gene expression in human  
 XX brains -  
 PT Example 4; SEQ ID NO: 28328; 650pp + Sequence Listing; English.  
 XX The present invention provides a number of single exon nucleic acid  
 CC probes which are derived from genomic sequences expressed in the human  
 CC brain. They can be used to measure gene expression in brain cell samples,  
 CC which may enable the diagnosis and improved treatment of nervous system  
 CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,  
 CC epilepsy and cancers. The present sequence is a protein encoded by one of  
 CC the probes of the invention.  
 XX Sequence 53 AA;  
 SQ Query Match 76.5%; Score 13; DB 22; Length 53;  
 Best Local Similarity 50.0%; Pred. No. 2e+03; 3; Indels 0; Gaps 0;  
 Matches 3; Conservative 0; Mismatches 0;  
 QY 2 EXXXLL 7  
 Db 32 EATALL 37  
 RESULT 27  
 AAM68600  
 ID AAM68600 standard; Protein; 53 AA.  
 XX AC AAM68600;  
 XX 06-NOV-2001 (first entry)  
 DT Human bone marrow expressed probe encoded protein SEQ ID NO: 28906.  
 DE Human; bone marrow expressed exon; gene expression analysis; probe;  
 XX microarray; cancer; leukaemia; lymphoma; myeloma.  
 KW Homo sapiens.  
 XX WO200157276-A2.  
 PN 09-AUG-2001.  
 PD 30-JAN-2001; 2001WO-US00668.  
 PF 04-FEB-2000; 2000US-0180312.  
 PR 26-MAY-2000; 2000US-0207456.  
 PR 30-JUN-2000; 2000US-0608408.  
 PR 03-AUG-2000; 2000US-0632366.  
 PR 21-SEP-2000; 2000US-0234687.  
 PR 27-SEP-2000; 2000US-0236359.  
 PR 04-OCT-2000; 2000GB-0024263.  
 XX (MOLE-) MOLECULAR DYNAMICS INC.  
 PA Penn SG, Hanzel DK, Chen W, Rank DR;  
 PI WPI; 2001-488899/53.  
 DR Single exon nucleic acid probes for analyzing gene expression in human  
 XX hearts -  
 PT Claim 15; SEQ ID No 22606; 530pp; English.  
 XX The present invention relates to single exon nucleic acid probes for  
 CC measuring human gene expression in a sample derived from human heart (see  
 CC ABA21535-ABA41305). The present sequence is a protein encoded by one such  
 CC probe. The probes may be used for predicting, measuring and displaying  
 CC gene expression in samples derived from the human heart via microarrays.  
 CC By measuring gene expression, the probes are useful for predicting, the  
 CC diagnosing, grading, staging, monitoring and prognosing diseases of the  
 CC human heart and vascular system e.g. cardiovascular disease,  
 CC hypertension, cardiac arrhythmias and congenital heart disease.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX Sequence 53 AA;  
 SQ Query Match 76.5%; Score 13; DB 22; Length 53;  
 Best Local Similarity 50.0%; Pred. No. 2e+03; 3; Indels 0; Gaps 0;  
 Matches 3; Conservative 0; Mismatches 0;  
 QY 2 EXXXLL 7  
 Db 32 EATALL 37  
 RESULT 26  
 AAM56223  
 ID AAM56223 standard; Protein; 53 AA.  
 XX AC AAM56223;  
 XX 05-NOV-2001 (first entry)  
 DT

RESULT 25  
 ABB20836  
 ID ABB20836 standard; Protein; 53 AA.  
 XX AC ABB20836;  
 XX 23-JAN-2002 (first entry)  
 DT Protein #2835 encoded by probe for measuring heart cell gene expression.  
 DE Human; gene expression; heart; microarray; vascular system;  
 KW cardiovascular disease; hypertension; cardiac arrhythmia;  
 KW congenital heart disease.  
 XX Homo sapiens.  
 XX WO200157274-A2.  
 PN 09-AUG-2001.  
 PD 30-JAN-2001; 2001WO-US00666.  
 PF 04-FEB-2000; 2000US-0180312.  
 PR 26-MAY-2000; 2000US-0207456.  
 PR 30-JUN-2000; 2000US-0608408.  
 PR 03-AUG-2000; 2000US-0632366.  
 PR 21-SEP-2000; 2000US-0234687.  
 PR 27-SEP-2000; 2000US-0236359.  
 PR 04-OCT-2000; 2000GB-0024263.  
 XX (MOLE-) MOLECULAR DYNAMICS INC.  
 PA Penn SG, Hanzel DK, Chen W, Rank DR;  
 PI WPI; 2001-488899/53.  
 DR Single exon nucleic acid probes for analyzing gene expression in human  
 XX hearts -  
 PT Claim 15; SEQ ID No 22606; 530pp; English.  
 XX The present invention relates to single exon nucleic acid probes for  
 CC measuring human gene expression in a sample derived from human heart (see  
 CC ABA21535-ABA41305). The present sequence is a protein encoded by one such  
 CC probe. The probes may be used for predicting, measuring and displaying  
 CC gene expression in samples derived from the human heart via microarrays.  
 CC By measuring gene expression, the probes are useful for predicting, the  
 CC diagnosing, grading, staging, monitoring and prognosing diseases of the  
 CC human heart and vascular system e.g. cardiovascular disease,  
 CC hypertension, cardiac arrhythmias and congenital heart disease.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX Sequence 53 AA;  
 SQ Query Match 76.5%; Score 13; DB 22; Length 53;  
 Best Local Similarity 50.0%; Pred. No. 2e+03; 3; Indels 0; Gaps 0;  
 Matches 3; Conservative 0; Mismatches 0;  
 QY 2 EXXXLL 7  
 Db 32 EATALL 37  
 RESULT 26  
 AAM56223  
 ID AAM56223 standard; Protein; 53 AA.  
 XX AC AAM56223;  
 XX 05-NOV-2001 (first entry)  
 DT

PR 04-FEB-2000; 2000US-0180312.  
 PR 26-MAY-2000; 2000US-0207456.  
 PR 30-JUN-2000; 2000US-0608408.  
 PR 03-AUG-2000; 2000US-0632366.  
 PR 21-SEP-2000; 2000US-0234687.  
 PR 27-SEP-2000; 2000US-0236359.  
 PR 04-OCT-2000; 2000GB-0024263.  
 XX (MOLE-) MOLECULAR DYNAMICS INC.  
 XX Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX WPI; 2001-488900/53.  
 XX Human genome-derived single exon nucleic acid probes useful for  
 XX analyzing gene expression in human bone marrow -  
 PS Example 4; SEQ ID NO: 28906; 658pp + Sequence Listing; English.  
 XX The present invention provides a number of single exon nucleic acid  
 CC probes which are derived from genomic sequences expressed in the human  
 CC bone marrow. They can be used to measure gene expression in bone marrow  
 CC samples, which may enable the improved diagnosis and treatment of cancers  
 CC such as lymphoma, leukaemia and myeloma. The present sequence is a  
 CC protein encoded by one of the probes of the invention.  
 XX Sequence 53 AA;  
 SQ Query Match 76.5%; Score 13; DB 22; Length 53;  
 Best Local Similarity 50.0%; Pred. No. 2e+03;  
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 QY 2 EXXXLL 7  
 DB 32 EATALL 37

RESULT 28  
 AAM16408  
 ID AAM16408 standard; Protein; 53 AA.  
 XX AC AAM16408;  
 XX DT 12-OCT-2001 (first entry)  
 XX DE Peptide #2842 encoded by probe for measuring cervical gene expression.  
 XX Probe; human; microarray; gene expression; cervical epithelial cell;  
 XX cervical cancer.  
 XX OS Homo sapiens.  
 XX PN WO200157278-A2.  
 XX PD 09-AUG-2001.  
 XX PP 30-JAN-2001; 2001WO-US00670.  
 XX 04-FEB-2000; 2000US-0180312.  
 XX 26-MAY-2000; 2000US-0207456.  
 XX 30-JUN-2000; 2000US-0608408.  
 XX 03-AUG-2000; 2000US-0632366.  
 XX 21-SEP-2000; 2000US-0234687.  
 XX 27-SEP-2000; 2000US-0236359.  
 XX 04-OCT-2000; 2000GB-0024263.  
 XX (MOLE-) MOLECULAR DYNAMICS INC.  
 XX Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX WPI; 2001-488901/53.  
 XX Human genome-derived single exon nucleic acid probes useful for

PT analyzing gene expression in human cervical epithelial cells -  
 XX Claim 27; SEQ ID No 21234; 487pp; English.  
 XX The present invention relates to human single exon nucleic acid probes  
 CC (SENP: see AAI10068-AAI28459). The present sequence is a peptide encoded  
 CC by one such probe. The SENPs are derived from human HeLa cells. The SENPs  
 CC can be used to produce a single exon microarray, which can be used for  
 CC measuring human gene expression in a sample derived from human cervical  
 CC epithelial cells. By measuring gene expression, the probes are therefore  
 CC useful in grading and/or staging of diseases of the cervix, notably  
 CC cervical cancer.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 XX at ftp.wipo.int/pub/published\_pct\_sequences.  
 SQ Sequence 53 AA;  
 Query Match 76.5%; Score 13; DB 22; Length 53;  
 Best Local Similarity 50.0%; Pred. No. 2e+03;  
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 QY 2 EXXXLL 7  
 DB 32 EATALL 37

RESULT 29  
 AAM28907  
 ID AAM28907 standard; Protein; 53 AA.  
 XX AC AAM28907;  
 XX DT 17-OCT-2001 (first entry)  
 XX DE Peptide #2944 encoded by probe for measuring placental gene expression.  
 XX Probe; microarray; human; placenta; antenatal diagnosis;  
 XX genetic disorder.  
 XX OS Homo sapiens.  
 XX PN WO200157272-A2.  
 XX PD 09-AUG-2001.  
 XX PP 30-JAN-2001; 2001WO-US00663.  
 XX 04-FEB-2000; 2000US-0180312.  
 XX 26-MAY-2000; 2000US-0207456.  
 XX 30-JUN-2000; 2000US-0608408.  
 XX 03-AUG-2000; 2000US-0632366.  
 XX 21-SEP-2000; 2000US-0234687.  
 XX 27-SEP-2000; 2000US-0236359.  
 XX 04-OCT-2000; 2000GB-0024263.  
 XX (MOLE-) MOLECULAR DYNAMICS INC.  
 XX Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX WPI; 2001-488997/53.  
 XX Human genome-derived single exon nucleic acid probes useful for  
 XX analyzing gene expression in human placenta -  
 PS Claim 27; SEQ ID No 29176; 654pp; English.  
 XX The present invention relates to single exon nucleic acid probes (SENP:  
 CC see AAI31315-AAI57546). The present sequence is a peptide encoded by one  
 CC such probe. The probes are useful for producing a microarray for  
 CC predicting, measuring and displaying gene expression in samples derived  
 CC from human placenta. The probes are useful for antenatal diagnosis of  
 CC human genetic disorders.

[illegible]



PR 08-SEP-2000; 2000US-0231414.  
 PR 08-SEP-2000; 2000US-0232080.  
 PR 08-SEP-2000; 2000US-0232081.  
 PR 12-SEP-2000; 2000US-0231968.  
 PR 14-SEP-2000; 2000US-0232397.  
 PR 14-SEP-2000; 2000US-0232398.  
 PR 14-SEP-2000; 2000US-0232399.  
 PR 14-SEP-2000; 2000US-0232400.  
 PR 14-SEP-2000; 2000US-0232401.  
 PR 14-SEP-2000; 2000US-0233063.  
 PR 14-SEP-2000; 2000US-0233064.  
 PR 14-SEP-2000; 2000US-0233065.  
 PR 21-SEP-2000; 2000US-0234223.  
 PR 21-SEP-2000; 2000US-0234274.  
 PR 25-SEP-2000; 2000US-0234997.  
 PR 25-SEP-2000; 2000US-0234998.  
 PR 26-SEP-2000; 2000US-0235484.  
 PR 27-SEP-2000; 2000US-0235834.  
 PR 27-SEP-2000; 2000US-0235836.  
 PR 29-SEP-2000; 2000US-0236327.  
 PR 29-SEP-2000; 2000US-0236367.  
 PR 29-SEP-2000; 2000US-0236368.  
 PR 29-SEP-2000; 2000US-0236369.  
 PR 29-SEP-2000; 2000US-0236370.  
 PR 02-OCT-2000; 2000US-0236802.  
 PR 02-OCT-2000; 2000US-0237037.  
 PR 02-OCT-2000; 2000US-0237038.  
 PR 02-OCT-2000; 2000US-0237039.  
 PR 02-OCT-2000; 2000US-0237040.  
 PR 13-OCT-2000; 2000US-0239935.  
 PR 13-OCT-2000; 2000US-0239937.  
 PR 20-OCT-2000; 2000US-0240960.  
 PR 20-OCT-2000; 2000US-0241221.  
 PR 20-OCT-2000; 2000US-0241225.  
 PR 20-OCT-2000; 2000US-0241785.  
 PR 20-OCT-2000; 2000US-0241786.  
 PR 20-OCT-2000; 2000US-0241787.  
 PR 20-OCT-2000; 2000US-0241808.  
 PR 20-OCT-2000; 2000US-0241809.  
 PR 01-NOV-2000; 2000US-0241826.  
 PR 08-NOV-2000; 2000US-0244617.  
 PR 08-NOV-2000; 2000US-0244674.  
 PR 08-NOV-2000; 2000US-0244675.  
 PR 08-NOV-2000; 2000US-0244676.  
 PR 08-NOV-2000; 2000US-0246477.  
 PR 08-NOV-2000; 2000US-0246478.  
 PR 08-NOV-2000; 2000US-0246523.  
 PR 08-NOV-2000; 2000US-0246524.  
 PR 08-NOV-2000; 2000US-0246525.  
 PR 08-NOV-2000; 2000US-0246526.  
 PR 08-NOV-2000; 2000US-0246527.  
 PR 08-NOV-2000; 2000US-0246528.  
 PR 08-NOV-2000; 2000US-0246532.  
 PR 08-NOV-2000; 2000US-0246609.  
 PR 08-NOV-2000; 2000US-0246610.  
 PR 08-NOV-2000; 2000US-0246611.  
 PR 08-NOV-2000; 2000US-0246613.  
 PR 17-NOV-2000; 2000US-0249207.  
 PR 17-NOV-2000; 2000US-0249208.  
 PR 17-NOV-2000; 2000US-0249209.  
 PR 17-NOV-2000; 2000US-0249210.  
 PR 17-NOV-2000; 2000US-0249211.  
 PR 17-NOV-2000; 2000US-0249212.  
 PR 17-NOV-2000; 2000US-0249213.  
 PR 17-NOV-2000; 2000US-0249214.  
 PR 17-NOV-2000; 2000US-0249215.  
 PR 17-NOV-2000; 2000US-0249216.  
 PR 17-NOV-2000; 2000US-0249217.  
 PR 17-NOV-2000; 2000US-0249218.  
 PR 17-NOV-2000; 2000US-0249244.  
 PR 17-NOV-2000; 2000US-0249245.  
 PR 17-NOV-2000; 2000US-0249264.  
 PR 17-NOV-2000; 2000US-0249265.  
 PR 17-NOV-2000; 2000US-0249297.

PR 17-NOV-2000; 2000US-0249299.  
 PR 17-NOV-2000; 2000US-0249300.  
 PR 01-DEC-2000; 2000US-0250160.  
 PR 01-DEC-2000; 2000US-0250391.  
 PR 05-DEC-2000; 2000US-0251030.  
 PR 05-DEC-2000; 2000US-0251988.  
 PR 05-DEC-2000; 2000US-0256719.  
 PR 06-DEC-2000; 2000US-0251479.  
 PR 08-DEC-2000; 2000US-0251856.  
 PR 08-DEC-2000; 2000US-0251868.  
 PR 08-DEC-2000; 2000US-0251869.  
 PR 08-DEC-2000; 2000US-0251989.  
 PR 11-DEC-2000; 2000US-0251990.  
 PR 11-DEC-2000; 2000US-0254097.  
 PR 05-JAN-2001; 2001US-0259678.  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 PA Rosen CA, Barash SC, Ruben SM;  
 FI WPI; 2001-483426/52.  
 XX N-PSDB; AAK58574.  
 DR Nucleic acids encoding human immune/hematopoietic antigen polypeptides,  
 DR useful for preventing, diagnosing and/or treating cancers and  
 DR metastasis -  
 PT Claim 11; SEQ ID NO 13386; 3071pp + Sequence Listing; English.  
 PT AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)  
 XX amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic  
 CC activity, and can be used in gene therapy and vaccine production. (I)  
 CC proteins and polynucleotides may be used in the prevention, diagnosis and  
 CC treatment of diseases associated with inappropriate (I) expression. For  
 CC example, they may be used to treat disorders associated with decreased  
 CC expression by rectifying mutations or deletions in a patient's genome  
 CC that affect the activity of (I) by expressing inactive proteins or to  
 CC supplement the patient's own production of (I). Additionally, (I)  
 CC polynucleotides may be used to produce the secreted (I), by inserting  
 CC the nucleic acids into a host cell and culturing the cell to express the  
 CC protein. (I) proteins and polynucleotides may be used to prevent,  
 CC diagnose and treat immune/haematopoietic-related diseases, especially  
 CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703  
 CC to AAK87694 represent human immune/haematopoietic antigen genomic  
 CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169  
 CC represent sequences used in the exemplification of the present invention.  
 XX  
 XX  
 SQ Sequence 64 AA;  
 Query Match 76.5%; Score 13; DB 22; Length 64;  
 Best Local Similarity 50.0%; Pred. No. 2.3e+03;  
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 Oy 2 EXXXLL 7  
 Db 46 ETSSL 51  
 RESULT 32  
 AAG58664  
 ID AAG58664 standard; Protein; 66 AA.  
 XX AC AAG58664;  
 XX 18-OCT-2000 (first entry)  
 DT Arabidopsis thaliana protein fragment SEQ ID NO: 75767.  
 DE Protein identification; signal transduction pathway; metabolic pathway;  
 XX hybridisation assay; genetic mapping; gene expression control; promoter;  
 KW termination sequence.  
 KW Arabidopsis thaliana.  
 XX OS

XX EP1033405-A2.  
PN 06-SEP-2000.  
PD 25-FEB-2000; 2000EP-0301439.  
XX 25-FEB-1999; 99US-0121825.  
XX 05-MAR-1999; 99US-0123180.  
XX 09-MAR-1999; 99US-0123548.  
XX 23-MAR-1999; 99US-0125788.  
XX 25-MAR-1999; 99US-0126264.  
XX 29-MAR-1999; 99US-0126785.  
XX 01-APR-1999; 99US-0127462.  
XX 06-APR-1999; 99US-0128234.  
XX 08-APR-1999; 99US-0128714.  
XX 16-APR-1999; 99US-0129845.  
XX 19-APR-1999; 99US-0130077.  
XX 21-APR-1999; 99US-0130449.  
XX 23-APR-1999; 99US-0130510.  
XX 28-APR-1999; 99US-0130891.  
XX 30-APR-1999; 99US-0131449.  
XX 30-APR-1999; 99US-0132048.  
XX 04-MAY-1999; 99US-0132407.  
XX 05-MAY-1999; 99US-0132484.  
XX 06-MAY-1999; 99US-0132485.  
XX 07-MAY-1999; 99US-0132487.  
XX 11-MAY-1999; 99US-0132863.  
XX 14-MAY-1999; 99US-0134256.  
XX 14-MAY-1999; 99US-0134219.  
XX 14-MAY-1999; 99US-0134219.  
XX 14-MAY-1999; 99US-0134221.  
XX 18-MAY-1999; 99US-0134370.  
XX 19-MAY-1999; 99US-0134768.  
XX 20-MAY-1999; 99US-0134941.  
XX 21-MAY-1999; 99US-0135124.  
XX 21-MAY-1999; 99US-0135353.  
XX 24-MAY-1999; 99US-0135629.  
XX 25-MAY-1999; 99US-0136021.  
XX 27-MAY-1999; 99US-0136392.  
XX 28-MAY-1999; 99US-0136782.  
XX 01-JUN-1999; 99US-0137222.  
XX 03-JUN-1999; 99US-0137528.  
XX 04-JUN-1999; 99US-0137503.  
XX 07-JUN-1999; 99US-0137724.  
XX 08-JUN-1999; 99US-0138094.  
XX 10-JUN-1999; 99US-0138540.  
XX 10-JUN-1999; 99US-0138840.  
XX 14-JUN-1999; 99US-0139119.  
XX 16-JUN-1999; 99US-0139452.  
XX 17-JUN-1999; 99US-0139453.  
XX 18-JUN-1999; 99US-0139492.  
XX 18-JUN-1999; 99US-0139454.  
XX 18-JUN-1999; 99US-0139455.  
XX 18-JUN-1999; 99US-0139456.  
XX 18-JUN-1999; 99US-0139457.  
XX 18-JUN-1999; 99US-0139458.  
XX 18-JUN-1999; 99US-0139459.  
XX 18-JUN-1999; 99US-0139460.  
XX 18-JUN-1999; 99US-0139461.  
XX 18-JUN-1999; 99US-0139462.  
XX 18-JUN-1999; 99US-0139463.  
XX 18-JUN-1999; 99US-0139750.  
XX 21-JUN-1999; 99US-0139763.  
XX 21-JUN-1999; 99US-0139817.  
XX 22-JUN-1999; 99US-0139899.  
XX 23-JUN-1999; 99US-0140353.  
XX 23-JUN-1999; 99US-0140354.  
XX 24-JUN-1999; 99US-0140695.  
XX 28-JUN-1999; 99US-0140823.  
XX 29-JUN-1999; 99US-0140991.  
XX 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.

PR 16-SEP-1999; 99US-0154039.  
 PR 20-SEP-1999; 99US-0154779.  
 PR 22-SEP-1999; 99US-0155139.  
 PR 23-SEP-1999; 99US-0155486.  
 PR 24-SEP-1999; 99US-0155659.  
 PR 28-SEP-1999; 99US-0156458.  
 PR 29-SEP-1999; 99US-0156596.  
 PR 04-OCT-1999; 99US-0157117.  
 PR 05-OCT-1999; 99US-0157753.  
 PR 06-OCT-1999; 99US-0157865.  
 PR 07-OCT-1999; 99US-0158029.  
 PR 08-OCT-1999; 99US-0158232.  
 PR 12-OCT-1999; 99US-0158369.  
 PR 13-OCT-1999; 99US-0159293.  
 PR 13-OCT-1999; 99US-0159294.  
 PR 13-OCT-1999; 99US-0159295.  
 PR 14-OCT-1999; 99US-0159329.  
 PR 14-OCT-1999; 99US-0159330.  
 PR 14-OCT-1999; 99US-0159331.  
 PR 14-OCT-1999; 99US-0159637.  
 PR 14-OCT-1999; 99US-0159638.  
 PR 18-OCT-1999; 99US-0159584.  
 PR 21-OCT-1999; 99US-0160741.  
 PR 21-OCT-1999; 99US-0160767.  
 PR 21-OCT-1999; 99US-0160768.  
 PR 21-OCT-1999; 99US-0160770.  
 PR 21-OCT-1999; 99US-0160814.  
 PR 21-OCT-1999; 99US-0160815.  
 PR 22-OCT-1999; 99US-0160980.  
 PR 22-OCT-1999; 99US-0160981.  
 PR 22-OCT-1999; 99US-0160981.  
 PR 25-OCT-1999; 99US-0161404.  
 PR 25-OCT-1999; 99US-0161405.  
 PR 25-OCT-1999; 99US-0161406.  
 PR 26-OCT-1999; 99US-0161359.  
 PR 26-OCT-1999; 99US-0161360.  
 PR 26-OCT-1999; 99US-0161361.  
 PR 28-OCT-1999; 99US-0161920.  
 PR 28-OCT-1999; 99US-0161992.  
 PR 28-OCT-1999; 99US-0161993.  
 PR 29-OCT-1999; 99US-0162142.

Query Match 76.5%; Score 13; DB 21; Length 66;  
 Best Local Similarity 50.0%; Pred. No. 2.4e+03;  
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qv. 2 EXXXLL 7  
 L 50 ETASLL 55

RESULT 33  
 AAUS3716  
 ID AAUS3716 standard; Protein; 66 AA.

XX AC AAUS3716;

DT 27-FEB-2002 (first entry)

DE Propionibacterium acnes immunogenic protein #14612.

XX SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;  
 KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;  
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;  
 KW dermatological; osteopathic; neuroprotectant.

XX Propionibacterium acnes.

XX WO200181581-A2.

XX 01-NOV-2001.

XX 20-APR-2001; 2001WO-US12865.

XX 21-APR-2000; 2000US-199047P.  
 PR 02-JUN-2000; 2000US-208841P.  
 PR 07-JUL-2000; 2000US-216747P.  
 XX (CORI-) CORIXA CORP.  
 XX Skelky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;  
 PI L'maisonneuve J, Zhang Y, Jen S, Carter D;  
 XX WPI; 2001-616774/71.  
 DR N-PSDB; AAS59562.

PT Propionibacterium acnes polypeptides and nucleic acids useful for  
 PT vaccinating against and diagnosing infections, especially useful for  
 PT treating acne vulgaris -

XX Example 1; SEQ ID No 14911; 1069pp; English.

XX Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic  
 CC polypeptides. The proteins and their associated DNA sequences are used in  
 CC the treatment, prevention and diagnosis of medical conditions caused by  
 CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,  
 CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.  
 CC P. acnes is also involved in infections of bone, joints and the central  
 CC nervous system, however it is particularly involved in the inflammatory  
 CC lesions associated with acne vulgaris. A method for detecting the  
 CC presence or absence of P. acnes in a patient comprises contacting a  
 CC sample with a binding agent that binds to the proteins of the invention  
 CC and determining the amount of bound protein in the sample. The  
 CC polypeptides may be used as antigens in the production of antibodies  
 CC specific for P. acnes proteins. These antibodies can be used to  
 CC downregulate expression and activity of P. acnes polypeptides and  
 CC therefore treat P. acnes infections. The antibodies may also be used as  
 CC diagnostic agents for determining P. acnes presence, for example, by  
 CC enzyme linked immunosorbent assay (ELISA).  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX SQ Sequence 66 AA;

Query Match 76.5%; Score 13; DB 22; Length 66;  
 Best Local Similarity 50.0%; Pred. No. 2.4e+03;  
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7  
 Db 6 ETATLL 11

RESULT 34  
 ABG28772  
 ID ABG28772 standard; Protein; 66 AA.

XX AC ABG28772;

DT 18-FEB-2002 (first entry)

XX DE Novel human diagnostic protein #28763.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 KW food supplement; medical imaging; diagnostic; genetic disorder.

XX Homo sapiens.

XX WO200175067-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US08631.

XX 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.  
 PA (HYSE-) HYSEQ INC.  
 XX Drmanac RT, Liu C, Tang YT;  
 XX WPI; 2001-639362/73.  
 DR N-PSDB; AAS92959.  
 XX  
 XX New isolated polynucleotide and encoded polypeptides, useful in  
 PT diagnostics, forensics, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity  
 XX  
 XX Claim 20; SEQ ID No 59131; 103pp; English.  
 XX  
 XX The invention relates to isolated polynucleotide (I) and  
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,  
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
 CC mapping and gene mapping, and in recombinant production of (II). The  
 CC polynucleotides are also used in diagnostics as expressed sequence tags  
 CC for identifying expressed genes. (I) is useful in gene therapy techniques  
 CC to restore normal activity of (II) or to treat disease states involving  
 CC quantitating a polypeptide in tissue, as molecular weight markers and as  
 CC a food supplement. (II) and its binding partners are useful in medical  
 CC imaging of sites expressing (II). (I) and (II) are useful for treating  
 CC disorders involving aberrant protein expression or biological activity.  
 CC The polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. ABG00010-ABG0377 represent novel human  
 CC diagnostic amino acid sequences of the invention.  
 CC Note: The sequence data for this patent did not appear in the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 SQ Sequence 66 AA;  
 Query Match 76.5%; Score 13; DB 22; Length 66;  
 Best Local Similarity 50.0%; Pred. NO. 2.4e+03;  
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 QY 2 EXXXLL 7  
 DB 32 ESSLL 37  
 ABG59264  
 ID ABG59264 standard; Peptide; 67 AA.  
 XX  
 XX ABG59264;  
 XX  
 XX 25-FEB-2003 (first entry)  
 DT  
 XX Human liver peptide, SEQ ID No 37912.  
 DE  
 XX Human; liver; cirrhosis; hyperlipoproteinaemia; hyperlipidaemia;  
 KW hypercholesterolaemia; coronary heart disease.  
 KW  
 XX Homo sapiens.  
 OS  
 XX WO200157273-A2.  
 FN  
 XX 09-AUG-2001.  
 PD  
 XX 30-JAN-2001; 2001WO-US00664.  
 PF  
 XX 04-FEB-2000; 2000US-0180312.  
 PR 26-MAY-2000; 2000US-0207456.  
 PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.  
 PR 21-SEP-2000; 2000US-0234687.  
 PR 27-SEP-2000; 2000US-0236359.  
 PR 04-OCT-2000; 2000GB-0024263.  
 XX  
 XX (MOLE-) MOLECULAR DYNAMICS INC.  
 PA Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX WPI; 2001-488898/53.  
 XX  
 XX Human genome-derived single exon nucleic acid probes useful for  
 PT analysing gene expression in human adult liver -  
 XX  
 XX Claim 27; SEQ ID No 37912; 658pp; English.  
 XX  
 XX The invention relates to a single exon nucleic acid probe (SENP) (I) for  
 CC measuring human gene expression in a sample derived from human adult  
 CC liver, comprising one of 13109 defined nucleotide sequences given in the  
 CC specification (or complements/ fragments). The probe hybridises at high  
 CC stringency to a nucleic acid molecule expressed in the human adult  
 CC liver. (I) may be used for predicting, measuring and displaying gene  
 CC expression in samples derived from human adult liver. The genes  
 CC identified may be involved in genetic liver diseases such as cirrhosis,  
 CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which  
 CC is associated with coronary heart disease. ABG47348-ABG59930 represent  
 CC human liver single exon encoded peptides of the invention.  
 CC Note: The sequence information for this patent does not appear in the  
 CC printed specification but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 SQ Sequence 67 AA;  
 Query Match 76.5%; Score 13; DB 22; Length 67;  
 Best Local Similarity 50.0%; Pred. NO. 2.4e+03;  
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 QY 2 EXXXLL 7  
 DB 35 ETSALL 40  
 RESULT 36  
 AAU46377  
 ID AAU46377 standard; Protein; 67 AA.  
 XX  
 XX AAU46377;  
 AC  
 XX 27-FEB-2002 (first entry)  
 DT  
 XX Propionibacterium acnes immunogenic protein #7273.  
 DE  
 XX SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;  
 KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;  
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;  
 KW dermatological; osteopathic; neuroprotectant.  
 XX  
 XX Propionibacterium acnes.  
 OS  
 XX WO200181581-A2.  
 FN  
 XX 01-NOV-2001.  
 PD  
 XX 20-APR-2001; 2001WO-US12865.  
 PF  
 XX 21-APR-2000; 2000US-199047P.  
 PR 02-JUN-2000; 2000US-208841P.  
 PR 07-JUL-2000; 2000US-216747P.  
 XX  
 XX (CORI-) CORIXA CORP.  
 PA  
 XX Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;  
 PI L'maisonneuve J, Zhang Y, Jen S, Carter D;

XX WPI; 2001-616774/71.  
 DR N-PSDB; AAS9531.  
 XX  
 PT Propionibacterium acnes polypeptides and nucleic acids useful for  
 PT vaccinating against and diagnosing infections, especially useful for  
 PT treating acne vulgaris -  
 XX  
 PS Example 1; SEQ ID NO 7572; 1069pp; English.  
 CC  
 CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic  
 CC polypeptides. The proteins and their associated DNA sequences are used in  
 CC the treatment, prevention and diagnosis of medical conditions caused by  
 CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,  
 CC pustulosis, hypertostia and osteomyelitis), uveitis and endophthalmitis.  
 CC P. acnes is also involved in infections of bone, joints and the central  
 CC nervous system, however it is particularly involved in the inflammatory  
 CC lesions associated with acne vulgaris. A method for detecting the  
 CC presence or absence of P. acnes in a patient comprises contacting a  
 CC sample with a binding agent that binds to the proteins of the invention  
 CC and determining the amount of bound protein in the sample. The  
 CC polypeptides may be used as antigens in the production of antibodies  
 CC specific for P. acnes proteins. These antibodies can be used to  
 CC downregulate expression and activity of P. acnes polypeptides and  
 CC therefore treat P. acnes infections. The antibodies may also be used as  
 CC diagnostic agents for determining P. acnes presence, for example, by  
 CC enzyme linked immunosorbent assay (ELISA).  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 XX Sequence 67 AA;  
 Query Match 76.5%; Score 13; DB 22; Length 67;  
 Best Local Similarity 50.0%; Pred. No. 2.4e+03;  
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 QY 2 EXXXLL 7  
 DB 61 EASALL 66  
 RESULT 37  
 ABB43884  
 ID ABB43884 standard; Peptide; 67 AA.  
 XX  
 XX ABB43884;  
 04-FEB-2002 (first entry)  
 DE Peptide #11390 encoded by human foetal liver single exon probe.  
 XX  
 XX Human; foetal liver; gene expression; single exon nucleic acid probe.  
 KW Homo sapiens.  
 OS  
 XX WO200157277-A2.  
 PN  
 XX 09-AUG-2001.  
 PD  
 XX 30-JAN-2001; 2001WO-US00669.  
 PF  
 XX 04-FEB-2000; 2000US-0180312.  
 PR 26-MAY-2000; 2000US-0207456.  
 PR 30-JUN-2000; 2000US-0608408.  
 PR 03-AUG-2000; 2000US-0632366.  
 PR 21-SEP-2000; 2000US-0234687.  
 PR 27-SEP-2000; 2000US-0236359.  
 PR 04-OCT-2000; 2000GB-0024263.  
 XX  
 XX (MOLE-) MOLECULAR DYNAMICS INC.  
 PA Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX  
 XX WPI; 2001-483446/52.  
 XX  
 XX Single exon nucleic acid probes for analyzing gene expression in human  
 PT brains -  
 PT  
 XX Example 4; SEQ ID NO: 36979; 650pp + Sequence Listing; English.  
 PS  
 XX The present invention provides a number of single exon nucleic acid  
 CC probes which are derived from genomic sequences expressed in the human  
 CC brain. They can be used to measure gene expression in brain cell samples,

XX WPI; 2001-483447/52.  
 DR  
 XX Human genome-derived single exon nucleic acid probes useful for  
 PT analyzing gene expression in human foetal liver -  
 XX  
 PS Claim 27; SEQ ID NO 36519; 639pp + sequence listing; English.  
 CC  
 CC The invention relates to a single exon nucleic acid probe for  
 CC measuring human gene expression in a sample derived from human foetal  
 CC liver. The single exon nucleic acid probes may be used for predicting,  
 CC measuring and displaying gene expression in samples derived from human  
 CC foetal liver. The present invention is a peptide encoded by a single exon  
 CC nucleic acid probe of the invention.  
 CC Note: The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 XX Sequence 67 AA;  
 Query Match 76.5%; Score 13; DB 22; Length 67;  
 Best Local Similarity 50.0%; Pred. No. 2.4e+03;  
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 QY 2 EXXXLL 7  
 DB 35 ETSALL 40  
 RESULT 38  
 AAM64874  
 ID AAM64874 standard; Protein; 67 AA.  
 XX  
 XX AAM64874;  
 AC  
 XX 05-NOV-2001 (first entry)  
 DT  
 XX Human brain expressed single exon probe encoded protein SEQ ID NO: 36979.  
 DE  
 XX Human; brain expressed exon; gene expression analysis; probe;  
 KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;  
 KW epilepsy; cancer.  
 XX  
 XX Homo sapiens.  
 OS  
 XX WO200157275-A2.  
 PN  
 XX 09-AUG-2001.  
 PD  
 XX 30-JAN-2001; 2001WO-US00667.  
 PF  
 XX 04-FEB-2000; 2000US-0180312.  
 PR 26-MAY-2000; 2000US-0207456.  
 PR 30-JUN-2000; 2000US-0608408.  
 PR 03-AUG-2000; 2000US-0632366.  
 PR 21-SEP-2000; 2000US-0234687.  
 PR 27-SEP-2000; 2000US-0236359.  
 PR 04-OCT-2000; 2000GB-0024263.  
 XX  
 XX (MOLE-) MOLECULAR DYNAMICS INC.  
 PA Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX  
 XX WPI; 2001-483446/52.  
 XX  
 XX Single exon nucleic acid probes for analyzing gene expression in human  
 PT brains -  
 PT  
 XX Example 4; SEQ ID NO: 36979; 650pp + Sequence Listing; English.  
 PS  
 XX The present invention provides a number of single exon nucleic acid  
 CC probes which are derived from genomic sequences expressed in the human  
 CC brain. They can be used to measure gene expression in brain cell samples,

CC which may enable the diagnosis and improved treatment of nervous system  
 CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,  
 CC epilepsy and cancers. The present sequence is a protein encoded by one of  
 CC the probes of the invention.

SQ Sequence 67 AA;

Query Match 76.5%; Score 13; DB 22; Length 67;  
 Best Local Similarity 50.0%; Pred. No. 2.4e+03;  
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 EXXXLL 7  
 | | |  
 DB 35 ETSALL 40

RESULT 39  
 AAM77610  
 ID AAM77610 standard; Protein; 67 AA.

XX AAM77610;

Dr 06-NOV-2001 (first entry)  
 XX Human bone marrow expressed probe encoded protein SEQ ID NO: 37916.  
 DE Human bone marrow expressed probe encoded protein SEQ ID NO: 37916.  
 XX Human; bone marrow expressed exon; gene expression analysis; probe;  
 KW microarray; cancer; leukaemia; lymphoma; myeloma.

OS Homo sapiens.

XX WO2000157276-A2.

PN 09-AUG-2001.

PD 30-JAN-2001; 2001WO-US00668.

PF 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0034263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

PI WPI; 2001-488900/53.

PT Human genome-derived single exon nucleic acid probes useful for  
 PT analyzing gene expression in human bone marrow -

XX Example 4; SEQ ID NO: 37916; 658pp + Sequence Listing; English.

XX The present invention provides a number of single exon nucleic acid  
 CC probes which are derived from genomic sequences expressed in the human  
 CC bone marrow. They can be used to measure gene expression in bone marrow  
 CC samples, which may enable the improved diagnosis and treatment of cancers  
 CC such as lymphoma, leukaemia and myeloma. The present sequence is a  
 CC protein encoded by one of the probes of the invention.

SQ Sequence 67 AA;

Query Match 76.5%; Score 13; DB 22; Length 67;  
 Best Local Similarity 50.0%; Pred. No. 2.4e+03;  
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 EXXXLL 7  
 | | |  
 DB 35 ETSALL 40

RESULT 40  
 AAM91468

ID AAM91468 standard; Protein; 67 AA.

XX AAM91468;

AC AAM91468;

DT 07-NOV-2001 (first entry)

XX Human immune/haematopoietic antigen SEQ ID NO:19061.

XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;  
 KW cytostatic; Gene therapy; vaccine; metastasis.

XX Homo sapiens.

XX WO200157182-A2.

XX 09-AUG-2001.

XX 17-JAN-2001; 2001WO-US01354.

XX 31-JAN-2000; 2000US-0179065.

PR 04-FEB-2000; 2000US-0180628.

PR 24-FEB-2000; 2000US-0184664.

PR 02-MAR-2000; 2000US-0186350.

PR 16-MAR-2000; 2000US-0189874.

PR 17-MAR-2000; 2000US-0190076.

PR 18-APR-2000; 2000US-0198123.

PR 19-MAY-2000; 2000US-0205515.

PR 07-JUN-2000; 2000US-0209467.

PR 28-JUN-2000; 2000US-0214886.

PR 30-JUN-2000; 2000US-0215135.

PR 07-JUL-2000; 2000US-0216647.

PR 07-JUL-2000; 2000US-0216880.

PR 11-JUL-2000; 2000US-0217487.

PR 11-JUL-2000; 2000US-0217496.

PR 14-JUL-2000; 2000US-0218290.

PR 26-JUL-2000; 2000US-0220963.

PR 26-JUL-2000; 2000US-0220964.

PR 14-AUG-2000; 2000US-0224518.

PR 14-AUG-2000; 2000US-0224519.

PR 14-AUG-2000; 2000US-0225213.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225266.

PR 14-AUG-2000; 2000US-0225267.

PR 14-AUG-2000; 2000US-0225268.

PR 14-AUG-2000; 2000US-0225270.

PR 14-AUG-2000; 2000US-0225447.

PR 14-AUG-2000; 2000US-0225757.

PR 14-AUG-2000; 2000US-0225758.

PR 18-AUG-2000; 2000US-0225759.

PR 22-AUG-2000; 2000US-0226681.

PR 22-AUG-2000; 2000US-0226868.

PR 22-AUG-2000; 2000US-0227182.

PR 23-AUG-2000; 2000US-0227009.

PR 30-AUG-2000; 2000US-0228924.

PR 01-SEP-2000; 2000US-0229287.

PR 01-SEP-2000; 2000US-0229343.

PR 01-SEP-2000; 2000US-0229344.

PR 01-SEP-2000; 2000US-0229345.

PR 05-SEP-2000; 2000US-0229509.

PR 05-SEP-2000; 2000US-0229513.

PR 06-SEP-2000; 2000US-0230437.

PR 06-SEP-2000; 2000US-0230438.

PR 08-SEP-2000; 2000US-0231242.

PR 08-SEP-2000; 2000US-0231243.

PR 08-SEP-2000; 2000US-0231244.

PR 08-SEP-2000; 2000US-0231413.

PR 08-SEP-2000; 2000US-0231414.

PR 08-SEP-2000; 2000US-0232080.

PR 08-SEP-2000; 2000US-0232081.

PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 21-SEP-2000; 2000US-0234597.  
PR 25-SEP-2000; 2000US-0234598.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240360.  
PR 20-OCT-2000; 2000US-0241121.  
PR 20-OCT-2000; 2000US-0241122.  
PR 20-OCT-2000; 2000US-0241185.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246533.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249219.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.

PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.  
XX (HUMA-) HUMAN GENOME SCI INC.  
PA Rosen CA, Barash SC, Ruben SM;  
FI WPI; 2001-483426/52.  
XX N-PSDB; AAK64249.  
DR Nucleic acids encoding human immune/hematopoietic antigen polypeptides,  
DR useful for preventing, diagnosing and/or treating cancers and  
DR metastasis -  
XX PT  
XX Claim 11; SEQ ID NO 19061; 3071pp + Sequence Listing; English.  
XX CC AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)  
CC amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic  
CC activity, and can be used in gene therapy and vaccine production. (I)  
CC proteins and polynucleotides may be used in the prevention, diagnosis and  
CC treatment of diseases associated with inappropriate (I) expression. For  
CC example, they may be used to treat disorders associated with decreased  
CC expression by rectifying mutations or deletions in a patient's genome  
CC that affect the activity of (I) by expressing inactive proteins or to  
CC supplement the patient's own production of (I). Additionally, (I)  
CC polynucleotides may be used to produce the secreted (I), by inserting  
CC the nucleic acids into a host cell and culturing the cell to express the  
CC protein. (I) proteins and polynucleotides may be used to prevent,  
CC diagnose and treat immune/haematopoietic-related diseases, especially  
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703  
CC to AAK87694 represent human immune/haematopoietic antigen genomic  
CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169  
CC represent sequences used in the exemplification of the present invention.  
XX SQ Sequence 67 AA;

Query Match 76.5%; Score 13; DB 22; Length 67;  
Best Local Similarity 50.0%; Pred. No. 2.4e+03;  
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7  
Db 3 ETSSLL 8

Search completed: November 5, 2003, 14:16:40  
Job time : 43 secs

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2003, 14:15:53 ; Search time 21 Seconds  
(without alignments)  
14.104 Million cell updates/sec

Title: US-09-620-840C-2

Perfect score: 17

Sequence: 1 XEXXXXL 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

ched: 328717 seqs, 42310859 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

Issued Patents AA:\*  
1: /cgn2\_6/prodata/1/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/prodata/1/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/prodata/1/iaa/6A\_COMB.pep.\*  
4: /cgn2\_6/prodata/1/iaa/6B\_COMB.pep.\*  
5: /cgn2\_6/prodata/1/iaa/PCTUS\_COMB.pep.\*  
6: /cgn2\_6/prodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	13	76.5	10	3	US-08-946-329A-38
2	13	76.5	19	2	US-09-174-060-21
3	13	76.5	19	3	US-08-338-382-21
4	13	76.5	41	1	US-08-122-520C-2
5	13	76.5	139	4	US-09-328-352-7327
6	13	76.5	144	3	US-08-946-329A-68
7	13	76.5	145	3	US-08-946-329A-69
8	13	76.5	150	4	US-09-252-991A-28955
9	13	76.5	156	4	US-09-904-615-67
10	13	76.5	158	2	US-08-933-750C-15
11	13	76.5	158	3	US-09-234-613-15
12	13	76.5	158	4	US-09-134-001C-3416
13	13	76.5	163	4	US-08-138-452A-103
14	13	76.5	189	3	US-08-779-764A-28
15	13	76.5	189	3	US-08-779-764A-29
16	13	76.5	189	3	US-08-779-764A-30
17	13	76.5	200	1	US-08-455-001-4
18	13	76.5	200	5	PCT-US95-11869-4
19	13	76.5	211	2	US-08-164-292B-2
20	13	76.5	211	3	US-08-845-623-2
21	13	76.5	211	3	US-08-815-927-2
22	13	76.5	211	4	US-09-103-330-2
23	13	76.5	211	4	US-09-435-242-2
24	13	76.5	223	4	US-03-328-352-6289
25	13	76.5	237	3	US-09-320-878-11
26	13	76.5	237	3	US-09-105-537-18
27	13	76.5	237	4	US-09-657-440-11

Sequence 29558, A  
Sequence 7, Appli  
Sequence 300, App  
Sequence 4, Appli  
Sequence 25028, A  
Sequence 19111, A  
Sequence 1109, Ap  
Sequence 797, App  
Sequence 7154, App  
Sequence 4, Appli  
Sequence 4, Appli  
Sequence 2, Appli  
Sequence 2, Appli  
Sequence 2, Appli  
Sequence 30112, A  
Sequence 2, Appli  
Sequence 2, Appli

28 13 76.5 237 4 US-09-252-991A-29558  
29 13 76.5 245 4 US-09-858-664A-7  
30 13 76.5 247 4 US-08-858-207A-300  
31 13 76.5 250 2 US-08-737-825-4  
32 13 76.5 258 4 US-09-252-991A-25028  
33 13 76.5 268 4 US-09-252-991A-19111  
34 13 76.5 286 4 US-09-198-452A-1109  
35 13 76.5 292 4 US-09-198-452A-797  
36 13 76.5 313 4 US-09-107-532A-7154  
37 13 76.5 337 4 US-09-641-806-4  
38 13 76.5 337 4 US-09-723-129-4  
39 13 76.5 337 4 US-08-722-862-4  
40 13 76.5 342 4 US-09-641-806-2  
41 13 76.5 342 4 US-09-723-129-2  
42 13 76.5 342 4 US-09-722-862-2  
43 13 76.5 350 4 US-09-252-991A-30112  
44 13 76.5 370 3 US-08-369-822C-2  
45 13 76.5 370 3 US-08-582-776C-2

#### ALIGNMENTS

RESULT 1  
US-08-946-329A-38  
; Sequence 38, Application US/08946329A  
; Patent No. 6057091  
; GENERAL INFORMATION:  
; APPLICANT: Beachy, Philip A.  
; APPLICANT: Porter, Jeffrey A.  
; TITLE OF INVENTION: NOVEL HEDGEHOG-DERIVED POLYPEPTIDES  
; NUMBER OF SEQUENCES: 109  
; CORRESPONDENCE ADDRESS:  
; ADDRESSER: Fish & Richardson P.C.  
; STREET: 4225 Executive Square, Suite 1400  
; CITY: La Jolla  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 92037  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: Windows 95  
; SOFTWARE: FASTSEQ for Windows Version 2.0b  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/946.329A  
; FILING DATE: 07-OCT-1997  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/061,323  
; FILING DATE: 07-OCT-1996  
; APPLICATION NUMBER: 08/729,743  
; FILING DATE: 10-JUL-1996  
; APPLICATION NUMBER: 08/567,357  
; FILING DATE: 04-DEC-1995  
; APPLICATION NUMBER: 08/349,498  
; FILING DATE: 02-DEC-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Haile, Lisa A.  
; REGISTRATION NUMBER: 38,347  
; REFERENCE/DOCKET NUMBER: 07265/140001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 619/678-5070  
; TELEFAX: 619/678-5099  
; INFORMATION FOR SEQ ID NO: 38:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 10 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-08-946-329A-38

Query Match 76.5%; Score 13; DB 3; Length 10;  
Best Local Similarity 50.0%; Pred. No. 1e+02;



Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
Qy 2 EXXXLL 7  
Db 4 ESTALL 9

RESULT 2  
US-09-174-060-21  
; Sequence 21, Application US/09174060  
; Patent No. 5989554  
; GENERAL INFORMATION:  
; APPLICANT: Knuth, Mark W  
; APPLICANT: Haak-Frendscho, Mary  
; APPLICANT: Shultz, John W  
; APPLICANT: Lesley, Scott A  
; APPLICANT: Villars, Catherine E  
; TITLE OF INVENTION: HIGH LEVEL EXPRESSION AND FACILE  
; TITLE OF INVENTION: PURIFICATION OF PROTEINS, PEPTIDES AND CONJUGATES FOR  
; TITLE OF INVENTION: IMMUNIZATION, PURIFICATION AND DETECTION APPLICATIONS  
; NUMBER OF SEQUENCES: 22  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Ross & Stevens, S.C.  
; STREET: 1 South Pinckney St.  
; CITY: Madison  
; STATE: WI  
; COUNTRY: USA  
; ZIP: 53701  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/174,060  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/338,382  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Sara, Charles S  
; REGISTRATION NUMBER: 30,492  
; REFERENCE/DOCKET NUMBER: 34506.024  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 608-257-5353  
; TELEFAX: 608-257-9175  
; INFORMATION FOR SEQ ID NO: 21:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 19 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: unknown  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-09-174-060-21

Query Match 76.5%; Score 13; DB 2; Length 19;  
Best Local Similarity 50.0%; Pred. No. 1.8e+02;  
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 EXXXLL 7  
Db 1 ETSSLL 6

RESULT 3  
US-08-338-382-21  
; Sequence 21, Application US/08338382  
; Patent No. 6069230  
; GENERAL INFORMATION:  
; APPLICANT: Knuth, Mark W  
; APPLICANT: Haak-Frendscho, Mary  
; APPLICANT: Shultz, John W

; APPLICANT: Lesley, Scott A  
; APPLICANT: Villars, Catherine E  
; TITLE OF INVENTION: HIGH LEVEL EXPRESSION AND FACILE  
; TITLE OF INVENTION: PURIFICATION OF PROTEINS, PEPTIDES AND CONJUGATES FOR  
; TITLE OF INVENTION: IMMUNIZATION, PURIFICATION AND DETECTION APPLICATIONS  
; NUMBER OF SEQUENCES: 22  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Ross & Stevens, S.C.  
; STREET: 1 South Pinckney St.  
; CITY: Madison  
; STATE: WI  
; COUNTRY: USA  
; ZIP: 53701  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/338,382  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Sara, Charles S  
; REGISTRATION NUMBER: 30,492  
; REFERENCE/DOCKET NUMBER: 34506.024  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 608-257-5353  
; TELEFAX: 608-257-9175  
; INFORMATION FOR SEQ ID NO: 21:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 19 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: unknown  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-338-382-21

Query Match 76.5%; Score 13; DB 3; Length 19;  
Best Local Similarity 50.0%; Pred. No. 1.8e+02;  
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 EXXXLL 7  
Db 1 ETSSLL 6

RESULT 4  
US-08-122-520C-2  
; Sequence 2, Application US/08122520C  
; Patent No. 5639663  
; GENERAL INFORMATION:  
; APPLICANT: Crosby, William L.  
; APPLICANT: Datla, Raju S.S.  
; APPLICANT: Hammerlindl, Joseph K.  
; APPLICANT: Selvaraj, Gopalan  
; TITLE OF INVENTION: BIFUNCTIONAL GENETIC MARKERS  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: J. Wayne Anderson  
; STREET: Montreal Rd.  
; CITY: Ottawa  
; STATE: ON  
; COUNTRY: CANADA  
; ZIP: K1A 0R6  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/122,520C  
; FILING DATE: 11/29/92

```
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Anderson, J. Wayne
REGISTRATION NUMBER: 28,158
REFERENCE/DOCKET NUMBER: 10013-1A
TELEPHONE: (613) 993-3899
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 41 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-122-520C-2

Query Match          76.5%; Score 13; DB 1; Length 41;
Best Local Similarity 50.0%; Pred. No. 3.7e+02;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

2 EXXXLL 7
14 ESTLL 19
Db

RESULT 5
US-09-328-352-7327
Sequence 7327, Application US/09328352
Patent No. 6562958
GENERAL INFORMATION:
APPLICANT: Gary L. Breton et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
FILE REFERENCE: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
CURRENT APPLICATION NUMBER: US/09/328,352
CURRENT FILING DATE: 1999-06-04
NUMBER OF SEQ ID NOS: 8252
SEQ ID NO 7327
LENGTH: 139
TYPE: PRT
ORGANISM: Acinetobacter baumannii
US-09-328-352-7327

Query Match          76.5%; Score 13; DB 4; Length 139;
Best Local Similarity 50.0%; Pred. No. 1.1e+03;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

2 EXXXLL 7
37 ETSSL 42
Db

RESULT 6
US-08-946-329A-68
Sequence 68, Application US/08946329A
Patent No. 6057091
GENERAL INFORMATION:
APPLICANT: Beachy, Philip A.
TITLE OF INVENTION: NOVEL HEDGEHOG-DERIVED POLYPEPTIDES
NUMBER OF SEQUENCES: 109
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 4225 Executive Square, Suite 1400
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows 95
SOFTWARE: FastSEQ for Windows Version 2.0b
CURRENT APPLICATION DATA:
FILING DATE: 07-OCT-1997
APPLICATION NUMBER: 60/061,323
FILING DATE: 07-OCT-1996
APPLICATION NUMBER: 08/729,743
FILING DATE: 10-JUL-1996
APPLICATION NUMBER: 08/567,357
FILING DATE: 04-DEC-1995
APPLICATION NUMBER: 08/349,498
FILING DATE: 02-DEC-1994
ATTORNEY/AGENT INFORMATION:
NAME: Halle, Lisa A.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 07265/140001
TELECOMMUNICATION INFORMATION:
```

```
APPLICATION NUMBER: US/08/946,329A
FILING DATE: 07-OCT-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/061,323
FILING DATE: 07-OCT-1996
APPLICATION NUMBER: 08/729,743
FILING DATE: 10-JUL-1996
APPLICATION NUMBER: 08/567,357
FILING DATE: 04-DEC-1995
APPLICATION NUMBER: 08/349,498
FILING DATE: 02-DEC-1994
ATTORNEY/AGENT INFORMATION:
NAME: Halle, Lisa A.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 07265/140001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619/678-5070
TELEFAX: 619/678-5099
INFORMATION FOR SEQ ID NO: 68:
SEQUENCE CHARACTERISTICS:
LENGTH: 144 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: Peptide
US-08-946-329A-68

Query Match          76.5%; Score 13; DB 3; Length 144;
Best Local Similarity 50.0%; Pred. No. 1.2e+03;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

2 EXXXLL 7
5 ESTALL 10
Db

RESULT 7
US-08-946-329A-69
Sequence 69, Application US/08946329A
Patent No. 6057091
GENERAL INFORMATION:
APPLICANT: Beachy, Philip A.
TITLE OF INVENTION: NOVEL HEDGEHOG-DERIVED POLYPEPTIDES
NUMBER OF SEQUENCES: 109
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 4225 Executive Square, Suite 1400
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows 95
SOFTWARE: FastSEQ for Windows Version 2.0b
CURRENT APPLICATION DATA:
FILING DATE: 07-OCT-1997
APPLICATION NUMBER: 60/061,323
FILING DATE: 07-OCT-1996
APPLICATION NUMBER: 08/729,743
FILING DATE: 10-JUL-1996
APPLICATION NUMBER: 08/567,357
FILING DATE: 04-DEC-1995
APPLICATION NUMBER: 08/349,498
FILING DATE: 02-DEC-1994
ATTORNEY/AGENT INFORMATION:
NAME: Halle, Lisa A.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 07265/140001
TELECOMMUNICATION INFORMATION:
```

TELEPHONE: 619/678-5070  
TELEFAX: 619/678-5099  
INFORMATION FOR SEQ ID NO: 69:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 145 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-946-329A-69

Query Match 76.5%; Score 13; DB 3; Length 145;  
Best Local Similarity 50.0%; Pred. No. 1.2e+03;  
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7  
|  
|  
|  
DB 5 ESTALL 10

RESULT 8  
US-09-252-991A-28955  
Sequence 28955, Application US/09252991A  
Patent No. 6551795

GENERAL INFORMATION:  
APPLICANT: Marc J. Rubenfield et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
FILE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: 107196.136  
CURRENT APPLICATION NUMBER: US/09/252,991A  
CURRENT FILING DATE: 1999-02-18  
PRIOR APPLICATION NUMBER: US 60/074,788  
PRIOR FILING DATE: 1998-02-18  
PRIOR APPLICATION NUMBER: US 60/094,190  
PRIOR FILING DATE: 1998-07-27  
NUMBER OF SEQ ID NOS: 33142  
SEQ ID NO 28955  
LENGTH: 150  
TYPE: PRT  
ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-28955

Query Match 76.5%; Score 13; DB 4; Length 150;  
Best Local Similarity 50.0%; Pred. No. 1.2e+03;  
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7  
|  
|  
|  
DB 36 ETTLL 41

RESULT 9  
US-09-904-615-67  
Sequence 67, Application US/09904615  
Patent No. 6566325  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: 49 Human Secreted Proteins  
FILE REFERENCE: P2032P1  
CURRENT APPLICATION NUMBER: US/09/904,615  
CURRENT FILING DATE: 2001-07-16  
PRIOR APPLICATION NUMBER: 09/511,554  
PRIOR FILING DATE: 2000-02-23  
PRIOR APPLICATION NUMBER: 60/097,917  
PRIOR FILING DATE: 1998-08-25  
PRIOR APPLICATION NUMBER: 60/098,634  
PRIOR FILING DATE: 1998-08-31  
NUMBER OF SEQ ID NOS: 170  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 67  
LENGTH: 156  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:

NAME/KEY: SITE  
LOCATION: (156)  
OTHER INFORMATION: Xaa equals stop translation  
US-09-904-615-67

Query Match 76.5%; Score 13; DB 4; Length 156;  
Best Local Similarity 50.0%; Pred. No. 1.3e+03;  
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7  
|  
|  
|  
DB 100 EAAALL 105

RESULT 10  
US-08-933-750C-15  
Sequence 15, Application US/08933750C  
Patent No. 5932442

GENERAL INFORMATION:  
APPLICANT: Lal, Preeti  
APPLICANT: Hillman, Jennifer L.  
APPLICANT: Bandman, Olga  
APPLICANT: Shah, Purvi  
APPLICANT: Au-Young, Janice  
APPLICANT: Yue, Henry  
APPLICANT: Guegler, Karl J.  
APPLICANT: Corley, Neil C.  
TITLE OF INVENTION: HUMAN REGULATORY MOLECULES  
NUMBER OF SEQUENCES: 98  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/933,750C  
FILING DATE: September 23, 1997  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0356 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166  
TELEX:  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 158 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: BRSTTUT01  
CLONE: 606742  
US-08-933-750C-15

Query Match 76.5%; Score 13; DB 2; Length 158;  
Best Local Similarity 50.0%; Pred. No. 1.3e+03;  
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7  
|  
|  
|

Db 74 EAAALL 79

## RESULT 11

US-09-234-613-15  
; Sequence 15, Application US/09234613  
; Patent No. 6132973  
; GENERAL INFORMATION:  
; APPLICANT: Lal, Preeti  
; APPLICANT: Hillman, Jennifer L.  
; APPLICANT: Bandman, Olga  
; APPLICANT: Shah, Purvi  
; APPLICANT: Au-Young, Janice  
; APPLICANT: Yue, Henry  
; APPLICANT: Guegler, Karl J.  
; APPLICANT: Corley, Neil C.  
; TITLE OF INVENTION: HUMAN REGULATORY MOLECULES  
; NUMBER OF SEQUENCES: 98  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/234,613  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/933,750  
; FILING DATE: September 23, 1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy J.  
; REGISTRATION NUMBER: 36,749  
; REFERENCE/DOCKET NUMBER: PF-0356 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-855-0555  
; TELEFAX: 415-845-4166  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 15:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 158 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: BRSTTUT01  
; CLONE: 606742  
US-09-234-613-15

Query Match 76.5%; Score 13; DB 3; Length 158;  
Best Local Similarity 50.0%; Pred. No. 1.3e+03;  
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7

Db 74 EAAALL 79

## RESULT 12

US-09-134-001C-3416  
; Sequence 3416, Application US/09134001C  
; Patent No. 6380370  
; GENERAL INFORMATION:  
; APPLICANT: Lynn Doucette-Stamm et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS

FILE REFERENCE: GTC-007  
; CURRENT APPLICATION NUMBER: US/09/134,001C  
; CURRENT FILING DATE: 1998-08-13  
; PRIOR APPLICATION NUMBER: US 60/064,964  
; PRIOR FILING DATE: 1997-11-08  
; PRIOR APPLICATION NUMBER: US 60/055,779  
; PRIOR FILING DATE: 1997-08-14  
; NUMBER OF SEQ ID NOS: 5674  
; SEQ ID NO 3416  
; LENGTH: 158  
; TYPE: PRT  
; ORGANISM: Staphylococcus epidermidis  
US-09-134-001C-3416

Query Match 76.5%; Score 13; DB 4; Length 158;  
Best Local Similarity 50.0%; Pred. No. 1.3e+03;  
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7

Db 147 ETTSL 152

## RESULT 13

US-09-198-452A-103  
; Sequence 103, Application US/09198452A  
; Patent No. 6559294  
; GENERAL INFORMATION:  
; APPLICANT: Griffais, R.  
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments  
; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention  
; TITLE OF INVENTION: and treatment of infection  
; FILE REFERENCE: 9710-003-999  
; CURRENT APPLICATION NUMBER: US/09/198,452A  
; CURRENT FILING DATE: 1998-11-24  
; NUMBER OF SEQ ID NOS: 6849  
; SEQ ID NO 103  
; LENGTH: 163  
; TYPE: PRT  
; ORGANISM: Chlamydia pneumoniae  
US-09-198-452A-103

Query Match 76.5%; Score 13; DB 4; Length 163;  
Best Local Similarity 50.0%; Pred. No. 1.3e+03;  
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7

Db 30 EAAALL 35

## RESULT 14

US-08-779-764A-28  
; Sequence 28, Application US/08779764A  
; Patent No. 6057094  
; GENERAL INFORMATION:  
; APPLICANT: de la Torre, Juan C.  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR SCREENING  
; TITLE OF INVENTION: OF HUMAN BORNA DISEASE VIRUS  
; NUMBER OF SEQUENCES: 65  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: THE SCRIPPS RESEARCH INSTITUTE  
; STREET: 10550 NO. 6057094th Torrey Pines Road, TPC-8  
; CITY: La Jolla  
; STATE: California  
; COUNTRY: United States  
; ZIP: 92037  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/779,764A  
FILING DATE: 16-DEC-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Fitting, Thomas  
REGISTRATION NUMBER: 34,163  
REFERENCE/DOCKET NUMBER: TSRI 465.0  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 784-2937  
TELEFAX: (619) 784-9399  
INFORMATION FOR SEQ ID NO: 28:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 189 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-779-764A-28

Query Match 76.5%; Score 13; DB 3; Length 189;  
Best Local Similarity 50.0%; Pred. No. 1.5e+03;  
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

2 EXXLL 7  
Db 111 ETATLL 116

RESULT 15  
US-08-779-764A-29  
Sequence 29, Application US/08/779764A  
Patent No. 6057094  
GENERAL INFORMATION:  
APPLICANT: de la Torre, Juan C.  
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR SCREENING  
TITLE OF INVENTION: OF HUMAN BORNA DISEASE VIRUS  
NUMBER OF SEQUENCES: 65  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: THE SCRIPPS RESEARCH INSTITUTE  
STREET: 10550 No. 6057094th Torrey Pines Road, TPC-8  
CITY: La Jolla  
STATE: California  
COUNTRY: United States  
ZIP: 92037

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/779,764A  
FILING DATE: 16-DEC-1996  
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:  
NAME: Fitting, Thomas  
REGISTRATION NUMBER: 34,163  
REFERENCE/DOCKET NUMBER: TSRI 465.0  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 784-2937  
TELEFAX: (619) 784-9399  
INFORMATION FOR SEQ ID NO: 29:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 189 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-779-764A-29

Query Match 76.5%; Score 13; DB 3; Length 189;  
Best Local Similarity 50.0%; Pred. No. 1.5e+03;  
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

2 EXXLL 7  
Qy 111

Db 111 ETATLL 116

RESULT 16  
US-08-779-764A-30  
Sequence 30, Application US/08/779764A  
Patent No. 6057094  
GENERAL INFORMATION:  
APPLICANT: de la Torre, Juan C.  
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR SCREENING  
TITLE OF INVENTION: OF HUMAN BORNA DISEASE VIRUS  
NUMBER OF SEQUENCES: 65  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: THE SCRIPPS RESEARCH INSTITUTE  
STREET: 10550 No. 6057094th Torrey Pines Road, TPC-8  
CITY: La Jolla  
STATE: California  
COUNTRY: United States  
ZIP: 92037  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/779,764A  
FILING DATE: 16-DEC-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Fitting, Thomas  
REGISTRATION NUMBER: 34,163  
REFERENCE/DOCKET NUMBER: TSRI 465.0  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 784-2937  
TELEFAX: (619) 784-9399  
INFORMATION FOR SEQ ID NO: 30:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 189 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-779-764A-30

Query Match 76.5%; Score 13; DB 3; Length 189;  
Best Local Similarity 50.0%; Pred. No. 1.5e+03;  
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

2 EXXLL 7  
Db 111 ETATLL 116

RESULT 17  
US-08-455-001-4  
Sequence 4, Application US/08455001  
Patent No. 5795734  
GENERAL INFORMATION:  
APPLICANT: Flanagan, John G.  
APPLICANT: Cheng, Hwai-Jong  
TITLE OF INVENTION: EPH Receptor Ligands, and Uses Related  
TITLE OF INVENTION: Thereto  
NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD  
STREET: 60 State Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: ASCII (text)  
CURRENT APPLICATION DATA:  
APPLICANT: US/08/455,001  
FILING DATE: 31 MAY 1995  
CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: Vincent, Matthew P.  
REGISTRATION NUMBER: 36,709  
REFERENCE/DOCKET NUMBER: HMI-011CP2  
TELEPHONE: (617) 227-7400  
TELEFAX: (617) 227-5941  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 200 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
98-455-001-4

Query Match 76.5%; Score 13; DB 1; Length 200;  
Best Local Similarity 50.0%; Pred. No. 1.6e+03;  
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 EXXXLL 7  
| | |  
Db 5 EAAALL 10

RESULT 18  
PCT-US95-11869-4  
; Sequence 4, Application PC/TUS9511869  
; GENERAL INFORMATION:  
; APPLICANT:  
; TITLE OF INVENTION: EPH Receptor Ligands, and Uses Related  
; TITLE OF INVENTION: Thereto  
; NUMBER OF SEQUENCES: 5  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE:  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/11869  
; FILING DATE: 19-SEP-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Vincent, Matthew P.  
; REGISTRATION NUMBER: 36,709  
; REFERENCE/DOCKET NUMBER: HMI-011CP2  
; TELEPHONE: (617) 227-7400  
; TELEFAX: (617) 227-5941  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 200 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
PCT-US95-11869-4

Query Match 76.5%; Score 13; DB 5; Length 200;  
Best Local Similarity 50.0%; Pred. No. 1.6e+03;  
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 EXXXLL 7  
| | |  
Db 5 EAAALL 10

RESULT 19  
US-08-164-292B-2  
; Sequence 2, Application US/08164292B  
; Patent No. 5820868

GENERAL INFORMATION:  
; APPLICANT: MITTAL, SURESH K.  
; APPLICANT: GRAHAM, FRANK L.  
; APPLICANT: PREVEC, LUDVIK  
; APPLICANT: BABIUK, LORNE A.  
; TITLE OF INVENTION: RECOMBINANT PROTEIN PRODUCTION IN BOVINE  
; TITLE OF INVENTION: ADENOVIRUS EXPRESSION VECTOR SYSTEM  
; NUMBER OF SEQUENCES: 34  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MORRISON & FOERSTER  
; STREET: 345 California Street  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94104-2675  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/164,292B  
; FILING DATE: 09-DEC-1993  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: GRACEY, NANCY J.  
; REGISTRATION NUMBER: 28,216  
; REFERENCE/DOCKET NUMBER: 29310-20021.00  
; TELEPHONE: (415) 677-7000  
; TELEFAX: (415) 677-7522  
; TELEX: 34-0154  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 211 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-164-292B-2

Query Match 76.5%; Score 13; DB 2; Length 211;  
Best Local Similarity 50.0%; Pred. No. 1.7e+03;  
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 EXXXLL 7  
| | |  
Db 70 EASSLL 75

RESULT 20  
US-08-845-623-2  
; Sequence 2, Application US/08845623A  
; Patent No. 6001591  
; GENERAL INFORMATION:  
; APPLICANT: BABIUK, LORNE A.  
; APPLICANT: TIKOO, SURESH K.  
; APPLICANT: REDDY, POLICE S.  
; TITLE OF INVENTION: BOVINE ADENOVIRUS 3 GENOME  
; FILE REFERENCE: 293102002120  
; CURRENT APPLICATION NUMBER: US/08/845,623A  
; CURRENT FILING DATE: 1997-04-25  
; EARLIER APPLICATION NUMBER: 08/164,294  
; EARLIER FILING DATE: 1993-12-09  
; NUMBER OF SEQ ID NOS: 34  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 211  
; TYPE: PRP  
; ORGANISM: Bovine adenovirus type 3  
US-08-845-623-2

Query Match 76.5%; Score 13; DB 3; Length 211;  
Best Local Similarity 50.0%; Pred. No. 1.7e+03;

Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 EXXXLL 7  
|  
Db 70 EASLL 75

## RESULT 21

US-08-815-927-2  
; Sequence 2, Application US/08815927  
; Patent No. 6086890  
; GENERAL INFORMATION:  
; APPLICANT: MITTAL, SURESH K.  
; APPLICANT: GRAHAM, FRANK L.  
; APPLICANT: PREVIC, LUDVIK  
; APPLICANT: BABIUK, LORNE A.  
; TITLE OF INVENTION: RECOMBINANT PROTEIN PRODUCTION IN BOVINE ADENOVIRUS EXPRESSION  
; FILE REFERENCE: 293102002101  
; CURRENT APPLICATION NUMBER: US/08/815,927  
; EARLIER FILING DATE: 1997-03-13  
; EARLIER FILING DATE: 1993-12-09  
; NUMBER OF SEQ ID NOS: 34  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 211  
; TYPE: PRT  
; ORGANISM: Bovine adenovirus type 3  
US-08-815-927-2

Query Match 76.5%; Score 13; DB 3; Length 211;  
Best Local Similarity 50.0%; Pred. No. 1.7e+03;  
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 EXXXLL 7  
|  
Db 70 EASLL 75

## RESULT 22

US-09-103-330-2  
; Sequence 2, Application US/09103330A  
; Patent No. 6319716  
; GENERAL INFORMATION:  
; APPLICANT: TIKOO, SURESH K.  
; APPLICANT: BABIUK, LORNE A.  
; APPLICANT: REDDY, POLICE S.  
; TITLE OF INVENTION: ISOLATION OF MUTANTS IN THE E3 REGION OF THE  
; FILE REFERENCE: 293102002121  
; CURRENT APPLICATION NUMBER: US/09/103,330A  
; CURRENT FILING DATE: 1998-06-23  
; EARLIER FILING DATE: 1997-06-23  
; EARLIER FILING DATE: 1997-06-23  
; EARLIER FILING DATE: 1993-12-09  
; NUMBER OF SEQ ID NOS: 40  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 211  
; TYPE: PRT  
; ORGANISM: Bovine adenovirus type 3  
US-09-103-330-2

Query Match 76.5%; Score 13; DB 4; Length 211;  
Best Local Similarity 50.0%; Pred. No. 1.7e+03;  
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 EXXXLL 7  
|  
Db 70 EASLL 75

RESULT 23  
US-09-435-242-2  
; Sequence 2, Application US/09435242  
; Patent No. 6379944  
; GENERAL INFORMATION:  
; APPLICANT: MITTAL, SURESH K.  
; APPLICANT: GRAHAM, FRANK L.  
; APPLICANT: PREVIC, LUDVIK  
; APPLICANT: BABIUK, LORNE A.  
; TITLE OF INVENTION: MAMMALIAN CELL LINES EXPRESSING BOVINE ADENOVIRUS FUNCTIONS  
; FILE REFERENCE: 293102002102  
; CURRENT APPLICATION NUMBER: US/09/435,242  
; CURRENT FILING DATE: 1999-11-05  
; EARLIER FILING DATE: 1997-03-13  
; EARLIER FILING DATE: 1997-03-13  
; EARLIER FILING DATE: 1993-12-09  
; NUMBER OF SEQ ID NOS: 34  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 211  
; TYPE: PRT  
; ORGANISM: Bovine adenovirus type 3  
US-09-435-242-2

Query Match 76.5%; Score 13; DB 4; Length 211;  
Best Local Similarity 50.0%; Pred. No. 1.7e+03;  
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 EXXXLL 7  
|  
Db 70 EASLL 75

## RESULT 24

US-09-328-352-6289  
; Sequence 6289, Application US/09328352  
; Patent No. 6562958  
; GENERAL INFORMATION:  
; APPLICANT: Gary L. Breton et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACT  
; FILE REFERENCE: GTC99-03PA  
; CURRENT APPLICATION NUMBER: US/09/328,352  
; CURRENT FILING DATE: 1999-06-04  
; NUMBER OF SEQ ID NOS: 8252  
; SEQ ID NO 6289  
; LENGTH: 223  
; TYPE: PRT  
; ORGANISM: Acinetobacter baumannii  
US-09-328-352-6289

Query Match 76.5%; Score 13; DB 4; Length 223;  
Best Local Similarity 50.0%; Pred. No. 1.7e+03;  
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 EXXXLL 7  
|  
Db 13 EASLL 18

## RESULT 25

US-09-320-878-11  
; Sequence 11, Application US/09320878A  
; Patent No. 6117659  
; GENERAL INFORMATION:  
; APPLICANT: ASHLEY, Gary  
; APPLICANT: BETLACH, Mary C.  
; APPLICANT: BETLACH, Mary C.  
; APPLICANT: MCDANIEL, Robert  
; APPLICANT: TANG, Li  
; TITLE OF INVENTION: RECOMBINANT NARBONOLIDE POLYKETIDE SYNTHASE

FILE REFERENCE: 300622002120  
CURRENT APPLICATION NUMBER: US/09/320,878A  
CURRENT FILING DATE: 1999-05-27  
EARLIER APPLICATION NUMBER: CIP OF 09/141,908  
EARLIER FILING DATE: 1998-08-28  
EARLIER APPLICATION NUMBER: CIP OF 09/073,538  
EARLIER FILING DATE: 1998-05-06  
EARLIER APPLICATION NUMBER: CIP OF 08/846,247  
EARLIER FILING DATE: 1997-04-30  
EARLIER APPLICATION NUMBER: 60/119,139  
EARLIER FILING DATE: 1999-02-08  
EARLIER APPLICATION NUMBER: 60/100,880  
EARLIER FILING DATE: 1998-09-22  
EARLIER APPLICATION NUMBER: 60/087,080  
EARLIER FILING DATE: 1998-05-28  
NUMBER OF SEQ ID NOS: 34  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 11  
LENGTH: 237  
TYPE: PRT  
ORGANISM: Streptomyces venezuelae  
US-09-320-878-11

Query Match 76.5%; Score 13; DB 3; Length 237;  
Best Local Similarity 50.0%; Pred. No. 1.8e+03;  
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7  
DB 38 EASSLL 43

## RESULT 26

US-09-105-537-18  
Sequence 18, Application US/09105537A  
Patent No. 6265202  
GENERAL INFORMATION:  
APPLICANT: Sherman, D.H.  
APPLICANT: Liu, H.  
APPLICANT: Xue, Y.  
APPLICANT: Zhao, L.  
TITLE OF INVENTION: DNA encoding methymycin and pikromycin  
FILE REFERENCE: 600.438US1  
CURRENT APPLICATION NUMBER: US/09/105,537A  
CURRENT FILING DATE: 1998-06-26  
NUMBER OF SEQ ID NOS: 43  
SOFTWARE: FastSeq for Windows Version 3.0  
Q ID NO 18  
LENGTH: 237  
TYPE: PRT  
ORGANISM: Streptomyces venezuelae  
US-09-105-537-18

Query Match 76.5%; Score 13; DB 3; Length 237;  
Best Local Similarity 50.0%; Pred. No. 1.8e+03;  
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7  
DB 38 EASSLL 43

## RESULT 27

US-09-657-440-11  
Sequence 11, Application US/09657440  
Patent No. 6509455  
GENERAL INFORMATION:  
APPLICANT: ASHLEY, Gary  
APPLICANT: BETLACH, Melanie C.  
APPLICANT: BETLACH, Mary C.  
APPLICANT: MCDANIEL, Robert  
APPLICANT: TANG, Li  
TITLE OF INVENTION: RECOMBINANT NARBONOLIDE POLYKETIDE SYNTHASE

FILE REFERENCE: 300622002120  
CURRENT APPLICATION NUMBER: US/09/657,440  
CURRENT FILING DATE: 2000-09-07  
PRIOR APPLICATION NUMBER: 09/320,878  
PRIOR FILING DATE: 1999-05-27  
PRIOR APPLICATION NUMBER: CIP OF 09/141,908  
PRIOR FILING DATE: 1998-08-28  
NUMBER OF SEQ ID NOS: 34  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 11  
LENGTH: 237  
TYPE: PRT  
ORGANISM: Streptomyces venezuelae  
US-09-657-440-11

Query Match 76.5%; Score 13; DB 4; Length 237;  
Best Local Similarity 50.0%; Pred. No. 1.8e+03;  
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7  
DB 38 EASSLL 43

## RESULT 28

US-09-252-991A-29558  
Sequence 29558, Application US/09252991A  
Patent No. 6551795  
GENERAL INFORMATION:  
APPLICANT: Marc J. Rubenfield et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
FILE REFERENCE: 107196.136  
CURRENT APPLICATION NUMBER: US/09/252,991A  
CURRENT FILING DATE: 1999-02-18  
PRIOR APPLICATION NUMBER: US 60/074,788  
PRIOR FILING DATE: 1998-02-18  
PRIOR APPLICATION NUMBER: US 60/094,190  
PRIOR FILING DATE: 1998-07-27  
NUMBER OF SEQ ID NOS: 33142  
SEQ ID NO 29558  
LENGTH: 237  
TYPE: PRT  
ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-29558

Query Match 76.5%; Score 13; DB 4; Length 237;  
Best Local Similarity 50.0%; Pred. No. 1.8e+03;  
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7  
DB 118 ESSSLL 123

## RESULT 29

US-09-858-664A-7  
Sequence 7, Application US/09858664A  
Patent No. 6482624  
GENERAL INFORMATION:  
APPLICANT: Wei, Ming-Hui, et al.  
TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC  
ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES  
TITLE OF INVENTION: THERBOF  
FILE REFERENCE: CLO00927-CIP  
CURRENT APPLICATION NUMBER: US/09/858,664A  
CURRENT FILING DATE: 2001-05-17  
PRIOR APPLICATION NUMBER: 09/711,134  
PRIOR FILING DATE: 2000-11-11  
NUMBER OF SEQ ID NOS: 33  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 7  
LENGTH: 245



TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-858-664A-7

Query Match 76.5%; Score 13; DB 4; Length 245;  
Best Local Similarity 50.0%; Pred. No. 1.9e+03;  
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7  
DB 48 EAATLL 53

## RESULT 30

US-08-858-207A-300  
Sequence 300, Application US/08858207A  
Patent No. 6348328

## GENERAL INFORMATION:

APPLICANT: Black, Michael  
APPLICANT: Hodgson, John  
APPLICANT: Knowles, David  
APPLICANT: Nicholas, Richard  
APPLICANT: Stodola, Robert  
TITLE OF INVENTION: NO. 6348328el Compounds  
NUMBER OF SEQUENCES: 552  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SmithKline Beecham Corporation  
STREET: 709 Swedeland Road  
CITY: King of Prussia  
STATE: PA  
COUNTRY: USA  
ZIP: 19406-0939

## COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/858.207A  
FILING DATE: 09-MAY-1997

## CLASSIFICATION: 435

## PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/017670

FILING DATE: 14-MAY-1996

ATTORNEY/AGENT INFORMATION:

NAME: Gimml, Edward R

REGISTRATION NUMBER: 38,891

REFERENCE/DOCKET NUMBER: P50475

TELECOMMUNICATION INFORMATION:

TELEPHONE: 610-270-4478

TELEFAX: 610-270-5090

TELEX:

INFORMATION FOR SEQ ID NO: 300:

SEQUENCE CHARACTERISTICS:

LENGTH: 247 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: No. 6348328e

US-08-858-207A-300

Query Match 76.5%; Score 13; DB 4; Length 247;  
Best Local Similarity 50.0%; Pred. No. 1.9e+03;  
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7  
DB 176 EAAALL 181

## RESULT 31

US-08-737-825-4

Sequence 4, Application US/08737825

Patent No. 5871922

## GENERAL INFORMATION:

APPLICANT: SALMOND, GEORGE PEACOCK COPELAND  
APPLICANT: MCGOWAN, SIMON JAMES  
APPLICANT: SEBATHIA, MOHAMMED  
APPLICANT: COX, ANTHONY RICHARD JOHN  
APPLICANT: HOLDEN, MATTHEW THOMAS GEOFFREY  
APPLICANT: PORTER, LAUREN ELIZABETH  
APPLICANT: BYCROFT, BARRIE WALSHAM  
APPLICANT: WILLIAMS, PAUL

ADDRESSEE: STEWART, GORDON SIDNEY ANDERSON BIRNIE

TITLE OF INVENTION: GENES INVOLVED IN THE BIOSYNTHETIC PATHWAY OF CARBAPENEM

NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESS:

ADDRESSEE: Birch, Stewart, Kolasch & Birch, LLP

STREET: P.O. Box 747

CITY: Falls Church

STATE: Virginia

COUNTRY: USA

ZIP: 22040-0747

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: MICROSOFT WORD97

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/737,825

FILING DATE: 03-JUN-1997

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: SVENSSON, LEONARD R.

REGISTRATION NUMBER: 30,330

REFERENCE/DOCKET NUMBER: 1009-0105P

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703) 205-8000

TELEFAX: (703) 205-8050

TELEX: 248345

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 250 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL: NO

ORIGINAL SOURCE:

ORGANISM: Erwinia carotovora

US-08-737-825-4

Query Match 76.5%; Score 13; DB 2; Length 250;

Best Local Similarity 50.0%; Pred. No. 1.9e+03;

Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7

DB 180 ESSALL 185

## RESULT 32

US-09-252-991A-25028

Sequence 25028, Application US/09252991A

Patent No. 6551795

## GENERAL INFORMATION:

APPLICANT: Marc J. Rubenfield et al.

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

FILE REFERENCE: 107196.136

CURRENT APPLICATION NUMBER: US/09/252,991A

CURRENT FILING DATE: 1999-02-18

PRIOR APPLICATION NUMBER: US 60/074,788

PRIOR FILING DATE: 1998-02-18

PRIOR APPLICATION NUMBER: US 60/094,190

PRIOR FILING DATE: 1998-07-27

NUMBER OF SEQ ID NOS: 33142  
SEQ ID NO 25028  
LENGTH: 258  
TYPE: PRT  
ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-25028

Query Match 76.5%; Score 13; DB 4; Length 258;  
Best Local Similarity 50.0%; Pred. No. 2e+03;  
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7  
DB 112 EATALL 117

RESULT 33  
US-09-252-991A-19111  
Sequence 19111, Application US/09252991A  
Patent No. 6551795

GENERAL INFORMATION:  
APPLICANT: Marc J. Rubenfield et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
FILE REFERENCE: 107196.136  
CURRENT APPLICATION NUMBER: US/09/252,991A  
PRIOR FILING DATE: 1999-02-18  
PRIOR APPLICATION NUMBER: US 60/074,788  
PRIOR FILING DATE: 1998-02-18  
PRIOR APPLICATION NUMBER: US 60/094,190  
PRIOR FILING DATE: 1998-07-27  
NUMBER OF SEQ ID NOS: 33142  
SEQ ID NO 19111  
LENGTH: 268  
TYPE: PRT  
ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-19111

Query Match 76.5%; Score 13; DB 4; Length 268;  
Best Local Similarity 50.0%; Pred. No. 2.1e+03;  
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7  
DB 257 EATALL 262

RESULT 34  
US-09-198-452A-1109  
Sequence 1109, Application US/09198452A  
Patent No. 6559294

GENERAL INFORMATION:  
APPLICANT: Griffais, R.  
TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments thereof and uses thereof, in particular for the diagnosis, prevention and treatment of infection  
FILE REFERENCE: 9710-003-999  
CURRENT APPLICATION NUMBER: US/09/198,452A  
PRIOR FILING DATE: 1998-11-24  
NUMBER OF SEQ ID NOS: 6849  
SEQ ID NO 1109  
LENGTH: 286  
TYPE: PRT  
ORGANISM: Chlamydia pneumoniae  
US-09-198-452A-1109

Query Match 76.5%; Score 13; DB 4; Length 286;  
Best Local Similarity 50.0%; Pred. No. 2.2e+03;  
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7  
DB 274 EATALL 279

RESULT 35  
US-09-198-452A-797  
Sequence 797, Application US/09198452A  
Patent No. 6559294

GENERAL INFORMATION:  
APPLICANT: Griffais, R.  
TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments thereof and uses thereof, in particular for the diagnosis, prevention and treatment of infection  
FILE REFERENCE: 9710-003-999  
CURRENT APPLICATION NUMBER: US/09/198,452A  
PRIOR FILING DATE: 1998-11-24  
NUMBER OF SEQ ID NOS: 6849  
SEQ ID NO 797  
LENGTH: 292  
TYPE: PRT  
ORGANISM: Chlamydia pneumoniae  
US-09-198-452A-797

Query Match 76.5%; Score 13; DB 4; Length 292;  
Best Local Similarity 50.0%; Pred. No. 2.2e+03;  
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7  
DB 107 EATALL 112

RESULT 36  
US-09-107-532A-7154  
Sequence 7154, Application US/09107532A  
Patent No. 6583275

GENERAL INFORMATION:  
APPLICANT: Lynn A. Doucette-Stamm and David Bush  
TITLE OF INVENTION: ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS  
NUMBER OF SEQUENCES: 7310  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: GENOME THERAPEUTICS CORPORATION  
STREET: 100 Beaver Street  
CITY: Waltham  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02354

COMPUTER READABLE FORM:  
MEDIUM TYPE: CD-ROM ISO9660  
COMPUTER: PC  
OPERATING SYSTEM: <Unknown>  
SOFTWARE: ASCII

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/107,532A  
FILING DATE: 30-Jun-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/085,598  
FILING DATE: 14 May 1998  
APPLICATION NUMBER: 60/051571  
FILING DATE: July 2, 1997

ATTORNEY/AGENT INFORMATION:  
NAME: Ariniello, Pamela Deneka  
REGISTRATION NUMBER: 40,489  
REFERENCE/DOCKET NUMBER: GTC-012  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (781)893-5007  
TELEFAX: (781)893-8277

INFORMATION FOR SEQ ID NO: 7154:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 313 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: YES

ORIGINAL SOURCE:  
ORGANISM: Enterococcus faecium  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (B) LOCATION 1...313  
SEQUENCE DESCRIPTION: SEQ ID NO: 7154:  
US-09-107-532A-7154

Query Match 76.5%; Score 13; DB 4; Length 313;  
Best Local Similarity 50.0%; Pred. No. 2.4e+03;  
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Oy 2 EXXXLL 7  
Db 208 ESSTLL 213

RESULT 37  
US-09-641-806-4  
Sequence 4, Application US/09641806  
Patent No. 6395527

GENERAL INFORMATION:  
APPLICANT: Beraud, Christophe  
APPLICANT: Freedman, Richard  
TITLE OF INVENTION: motor proteins and methods for  
FILE REFERENCE: 1034  
CURRENT APPLICATION NUMBER: US/09/641,806  
PRIOR FILING DATE: 2000-08-17  
NUMBER OF SEQ ID NOS: 4  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 4  
LENGTH: 337  
TYPE: PRT  
ORGANISM: Human  
US-09-641-806-4

Query Match 76.5%; Score 13; DB 4; Length 337;  
Best Local Similarity 50.0%; Pred. No. 2.5e+03;  
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Oy 2 EXXXLL 7  
Db 95 EATSL 100

RESULT 38  
US-09-723-129-4  
Sequence 4, Application US/09723129  
Patent No. 6551787

GENERAL INFORMATION:  
APPLICANT: Beraud, Christophe  
APPLICANT: Freedman, Richard  
TITLE OF INVENTION: motor proteins and methods for  
FILE REFERENCE: 1034  
CURRENT APPLICATION NUMBER: US/09/723,129  
PRIOR FILING DATE: 2000-11-27  
CURRENT APPLICATION NUMBER: 09/641,806  
PRIOR FILING DATE: 2000-08-17  
NUMBER OF SEQ ID NOS: 4  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 4  
LENGTH: 337  
TYPE: PRT  
ORGANISM: Human  
US-09-723-129-4

Query Match 76.5%; Score 13; DB 4; Length 337;  
Best Local Similarity 50.0%; Pred. No. 2.5e+03;  
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Oy 2 EXXXLL 7

Db 95 EATSL 100

RESULT 39  
US-09-722-862-4  
Sequence 4, Application US/09722862  
Patent No. 6562610  
GENERAL INFORMATION:  
APPLICANT: Beraud, Christophe  
APPLICANT: Freedman, Richard  
TITLE OF INVENTION: motor proteins and methods for  
FILE REFERENCE: 1034  
CURRENT APPLICATION NUMBER: US/09/722,862  
PRIOR FILING DATE: 2000-11-27  
CURRENT APPLICATION NUMBER: 09/641,806  
PRIOR FILING DATE: 2000-08-17  
NUMBER OF SEQ ID NOS: 4  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 4  
LENGTH: 337  
TYPE: PRT  
ORGANISM: Human  
US-09-722-862-4

Query Match 76.5%; Score 13; DB 4; Length 337;  
Best Local Similarity 50.0%; Pred. No. 2.5e+03;  
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Oy 2 EXXXLL 7  
Db 95 EATSL 100

RESULT 40  
US-09-641-806-2  
Sequence 2, Application US/09641806  
Patent No. 6395527  
GENERAL INFORMATION:  
APPLICANT: Beraud, Christophe  
APPLICANT: Freedman, Richard  
TITLE OF INVENTION: motor proteins and methods for  
FILE REFERENCE: 1034  
CURRENT APPLICATION NUMBER: US/09/641,806  
CURRENT FILING DATE: 2000-08-17  
NUMBER OF SEQ ID NOS: 4  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 2  
LENGTH: 342  
TYPE: PRT  
ORGANISM: Human  
US-09-641-806-2

Query Match 76.5%; Score 13; DB 4; Length 342;  
Best Local Similarity 50.0%; Pred. No. 2.6e+03;  
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Oy 2 EXXXLL 7  
Db 96 EATSL 101

Search completed: November 5, 2003, 14:18:58  
Job time : 22 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2003, 14:17:59 : Search time 28 Seconds  
(without alignments)  
42.937 Million cell updates/sec

Title: US-09-620-840C-2

Perfect score: 17

Sequence: 1 EXXXLL 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Indexed: 644079 seqs, 171749292 residues

Total number of hits satisfying chosen parameters: 644079

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	13	76.5	7	10	US-09-211-424-20
2	13	76.5	10	11	US-09-845-731-1
3	13	76.5	10	11	US-09-572-404B-2009
4	13	76.5	11	10	US-09-845-715-1
5	13	76.5	12	10	US-09-846-349-1
6	13	76.5	14	10	US-09-845-730-1
7	13	76.5	16	10	US-09-846-345-1
8	13	76.5	17	10	US-09-846-346-1
9	13	76.5	53	9	US-09-864-761-36134
10	13	76.5	67	12	US-09-864-761-46113
11	13	76.5	74	12	US-10-029-386-29854
12	13	76.5	82	9	US-09-864-761-48608
13	13	76.5	88	9	US-09-864-761-48010
14	13	76.5	94	15	US-10-106-698-5724
15	13	76.5	98	9	US-09-864-761-37529

16	13	76.5	138	10	US-09-738-626-3698	Sequence 3698, Ap
17	13	76.5	144	12	US-10-138-473-30	Sequence 30, Appl
18	13	76.5	144	16	US-10-080-170-94	Sequence 94, Appl
19	13	76.5	144	16	US-10-080-170-518	Sequence 518, Appl
20	13	76.5	156	9	US-09-739-254-67	Sequence 67, Appl
21	13	76.5	156	9	US-09-304-615-67	Sequence 67, Appl
22	13	76.5	156	12	US-10-055-098-67	Sequence 67, Appl
23	13	76.5	156	15	US-10-054-988-67	Sequence 67, Appl
24	13	76.5	158	9	US-09-840-787-15	Sequence 15, Appl
25	13	76.5	205	11	US-09-764-891-4552	Sequence 4552, Ap
26	13	76.5	211	14	US-10-046-938-2	Sequence 2, Appl1
27	13	76.5	224	12	US-10-029-386-32343	Sequence 32343, A
28	13	76.5	230	10	US-09-893-737-184	Sequence 184, App
29	13	76.5	234	11	US-09-764-891-2999	Sequence 2999, Ap
30	13	76.5	237	10	US-09-861-289-18	Sequence 18, Appl
31	13	76.5	237	11	US-09-860-846-18	Sequence 18, Appl
32	13	76.5	237	11	US-09-988-384B-18	Sequence 18, Appl
33	13	76.5	237	11	US-09-836-821-18	Sequence 18, Appl
34	13	76.5	237	11	US-09-793-708-11	Sequence 11, Appl
35	13	76.5	237	12	US-10-160-539-11	Sequence 11, Appl
36	13	76.5	237	12	US-10-271-889-18	Sequence 18, Appl
37	13	76.5	245	9	US-09-858-664A-7	Sequence 7, Appl1
38	13	76.5	247	9	US-09-815-242-13461	Sequence 13461, A
39	13	76.5	247	9	US-09-815-242-13622	Sequence 13622, A
40	13	76.5	255	14	US-10-024-632-13	Sequence 13, Appl
41	13	76.5	271	12	US-09-769-744A-86	Sequence 86, Appl
42	13	76.5	281	15	US-10-295-403-46	Sequence 46, Appl
43	13	76.5	281	15	US-10-156-761-12667	Sequence 12667, A
44	13	76.5	283	9	US-09-925-302-693	Sequence 693, App
45	13	76.5	301	10	US-09-764-868-660	Sequence 660, App

#### ALIGNMENTS

##### RESULT 1

US-09-211-424-20  
; Sequence 20, Application US/09211424A  
; Patent No. US20020177231A1  
; GENERAL INFORMATION:  
; APPLICANT: Staunton et al.  
; TITLE OF INVENTION: Leupaxin Materials and Methods  
; FILE REFERENCE:  
; CURRENT APPLICATION NUMBER: US/09/211,424A  
; CURRENT FILING DATE: 1998-12-15  
; NUMBER OF SEQ ID NOS: 38  
; SEQ ID NO 20  
; LENGTH: 7  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: modified human  
; OTHER INFORMATION: sequence  
US-09-211-424-20

Query Match 76.5%; Score 13; DB 10; Length 7;  
Best Local Similarity 50.0%; Pred. No. 5.7e+05;  
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 EXXXLL 7

Db 2 EAAALL 7

##### RESULT 2

US-09-845-731-1  
; Sequence 1, Application US/09845731  
; Publication No. US20030004307A1  
; GENERAL INFORMATION:  
; APPLICANT: Jackowski, George  
; TITLE OF INVENTION: Biopolymer Marker Indicative Of Disease State Having A Molecular Weight Of 1211 Daltons  
; FILE REFERENCE: 2132.029

; CURRENT APPLICATION NUMBER: US/09/845,731  
; CURRENT FILING DATE: 2001-04-30  
; NUMBER OF SEQ ID NOS: 1  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1  
; TYPE: PRT  
; LENGTH: 10  
; ORGANISM: Homo sapiens  
US-09-845-731-1

Query Match 76.5%; Score 13; DB 11; Length 10;  
Best Local Similarity 50.0%; Pred. No. 2.5e+02;  
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7  
DB 5 ESASLL 10

## RESULT 3

US-09-572-404B-2009  
Sequence 2009, Application US/09572404B  
Publication No. US20030078374A1  
GENERAL INFORMATION:

; APPLICANT: Proteom Ltd  
; TITLE OF INVENTION: Complementary peptide ligands from the human genome  
; FILE REFERENCE: Human patent  
; CURRENT APPLICATION NUMBER: US/09/572,404B  
; CURRENT FILING DATE: 2000-05-17  
; NUMBER OF SEQ ID NOS: 4203  
; SOFTWARE: ProtPatent version 1.0  
; SEQ ID NO 2009  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Homo Sapiens

; FEATURE:  
; OTHER INFORMATION: sequence located in PDE6B OR PDEB at 54-63 and may interact with

US-09-572-404B-2009

Query Match 76.5%; Score 13; DB 11; Length 10;  
Best Local Similarity 50.0%; Pred. No. 2.5e+02;  
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7  
DB 2 ESTALL 7

## RESULT 4

US-09-845-715-1  
Sequence 1, Application US/09845715  
Patent No. US20020161184A1  
GENERAL INFORMATION:

; APPLICANT: Jackowski, George  
; TITLE OF INVENTION: BIOPOLYMER MARKER INDICATIVE OF DISEASE STATE HAVING A MOLECULAR  
; FILE REFERENCE: 2132.030  
; CURRENT APPLICATION NUMBER: US/09/845,715  
; CURRENT FILING DATE: 2001-04-30  
; NUMBER OF SEQ ID NOS: 1  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1  
; LENGTH: 11  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-845-715-1

Query Match 76.5%; Score 13; DB 10; Length 11;  
Best Local Similarity 50.0%; Pred. No. 2.8e+02;  
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7

DB 6 ESASLL 11

## RESULT 5

US-09-846-349-1  
Sequence 1, Application US/09846349  
Patent No. US20020161186A1  
GENERAL INFORMATION:

; APPLICANT: Jackowski, George  
; TITLE OF INVENTION: BIOPOLYMER MARKER INDICATIVE OF DISEASE STATE HAVING A MOLECULAR  
; FILE REFERENCE: 2132.034  
; CURRENT APPLICATION NUMBER: US/09/846,349  
; CURRENT FILING DATE: 2001-04-30  
; NUMBER OF SEQ ID NOS: 1  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1  
; LENGTH: 12  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-846-349-1

Query Match 76.5%; Score 13; DB 10; Length 12;  
Best Local Similarity 50.0%; Pred. No. 3e+02;  
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7  
DB 7 ESASLL 12

## RESULT 6

US-09-845-730-1  
Sequence 1, Application US/09845730  
Patent No. US20020169278A1  
GENERAL INFORMATION:

; APPLICANT: Jackowski, George  
; TITLE OF INVENTION: BIOPOLYMER MARKER INDICATIVE OF DISEASE STATE HAVING A MOLECULAR  
; FILE REFERENCE: 2132.042  
; CURRENT APPLICATION NUMBER: US/09/845,730  
; CURRENT FILING DATE: 2001-04-30  
; NUMBER OF SEQ ID NOS: 1  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1  
; LENGTH: 14  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-845-730-1

Query Match 76.5%; Score 13; DB 10; Length 14;  
Best Local Similarity 50.0%; Pred. No. 3.5e+02;  
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7  
DB 9 ESASLL 14

## RESULT 7

US-09-846-345-1  
Sequence 1, Application US/09846345  
Patent No. US20020161182A1  
GENERAL INFORMATION:

; APPLICANT: Jackowski, George  
; TITLE OF INVENTION: BIOPOLYMER MARKER INDICATIVE OF DISEASE STATE HAVING A MOLECULAR  
; FILE REFERENCE: 2132.045  
; CURRENT APPLICATION NUMBER: US/09/846,345  
; CURRENT FILING DATE: 2001-04-30  
; NUMBER OF SEQ ID NOS: 1  
; SOFTWARE: PatentIn version 3.1

```

; SEQ ID NO 1
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-846-345-1

Query Match
Best Local Similarity 76.5%; Score 13; DB 10; Length 16;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7
DB 11 ESASLL 16

RESULT 8
US-09-846-346-1
; Sequence 1, Application US/09846346
; Patent No. US20020160532A1
; GENERAL INFORMATION:
; APPLICANT: Jackowski, George
; TITLE OF INVENTION: BIOPOLYMER MARKER INDICATIVE OF DISEASE STATE HAVING A MOLECULAR
; FILE REFERENCE: 2132.013
; CURRENT APPLICATION NUMBER: US/09/846,346
; CURRENT FILING DATE: 2001-04-30
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-846-346-1

Query Match
Best Local Similarity 76.5%; Score 13; DB 10; Length 17;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7
DB 11 ESASLL 16

RESULT 9
US-09-864-761-36134
; Sequence 36134, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aomic-a-x-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30

```

```

; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 36134
; LENGTH: 53
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL121763.5
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 0.98
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.99
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.94
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.2
; OTHER INFORMATION: EST HUMAN HIT: AL039044.1, EVALUATE 3.00e-08
; OTHER INFORMATION: SWISSPROT HIT: O51596, EVALUATE 5.90e-00
US-09-864-761-36134

Query Match
Best Local Similarity 76.5%; Score 13; DB 9; Length 53;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7
DB 32 EATALL 37

RESULT 10
US-09-864-761-46113
; Sequence 46113, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aomic-a-x-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359

```

;; PRIOR FILING DATE: 2000-09-27  
;; PRIOR APPLICATION NUMBER: PCT/US01/00666  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00667  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00664  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00669  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00665  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00668  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00663  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00662  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00661  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00670  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: US 60/234,687  
;; PRIOR FILING DATE: 2000-09-21  
;; PRIOR APPLICATION NUMBER: US 09/608,408  
;; PRIOR FILING DATE: 2000-06-30  
;; PRIOR APPLICATION NUMBER: US 09/774,203  
;; PRIOR FILING DATE: 2001-01-29  
;; NUMBER OF SEQ ID NOS: 49117  
;; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
;; SEQ ID NO 46113  
;; LENGTH: 67  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
;;  
;; OTHER INFORMATION: MAP TO AC013528.3  
;; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.4  
;; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.5  
;; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.6  
;; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.5  
;; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.3  
;; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.4  
;; OTHER INFORMATION: EST HUMAN HIT: AM009639.1, EVALUE 6.50e-01  
;; OTHER INFORMATION: SWISSPROT HIT: P08648, EVALUE 1.00e-01  
US-09-864-761-46113

Query Match 76.5%; Score 13; DB 9; Length 67;  
Best Local Similarity 50.0%; Pred. No. 1.5e+03;  
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
DB 2 EXXXLL 7  
35 ETSALL 40

RESULT 11  
US-10-029-386-29854  
;; Sequence 29854, Application US/10029386  
;; Publication No. US20030194704A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Penn, Sharron G.  
;; APPLICANT: Hanzel, David K.  
;; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
;; FILE REFERENCE: AEOMICA-X-2  
;; CURRENT APPLICATION NUMBER: US/10/029,386  
;; CURRENT FILING DATE: 2001-12-20  
;; NUMBER OF SEQ ID NOS: 34288  
;; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
;; SEQ ID NO 29854  
;; LENGTH: 74  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens

;; FEATURE:  
;; OTHER INFORMATION: MAP TO CHR17.1  
;; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.3  
;; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.73  
;; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 0.75  
;; OTHER INFORMATION: SWISSPROT HIT: Q99536, EVALUE 4.00e-35  
US-10-029-386-29854  
Query Match 76.5%; Score 13; DB 12; Length 74;  
Best Local Similarity 50.0%; Pred. No. 1.6e+03;  
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
DB 2 EXXXLL 7  
37 EAAALL 42  
RESULT 12  
US-09-864-761-48608  
;; Sequence 48608, Application US/09864761  
;; Patent No. US20020048763A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Penn, Sharron G.  
;; APPLICANT: Rank, David R.  
;; APPLICANT: Hanzel, David K.  
;; APPLICANT: Chen, Wensheng  
;; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL  
;; FILE REFERENCE: AEOMICA-X-1  
;; CURRENT APPLICATION NUMBER: US/09/864,761  
;; CURRENT FILING DATE: 2001-05-23  
;; PRIOR APPLICATION NUMBER: US 60/180,312  
;; PRIOR FILING DATE: 2000-02-04  
;; PRIOR APPLICATION NUMBER: US 60/207,456  
;; PRIOR FILING DATE: 2000-05-26  
;; PRIOR APPLICATION NUMBER: US 09/632,366  
;; PRIOR FILING DATE: 2000-08-03  
;; PRIOR APPLICATION NUMBER: GB 24263.6  
;; PRIOR FILING DATE: 2000-10-04  
;; PRIOR APPLICATION NUMBER: US 60/236,359  
;; PRIOR FILING DATE: 2000-09-27  
;; PRIOR APPLICATION NUMBER: PCT/US01/00666  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00667  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00664  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00669  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00665  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00668  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00663  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00662  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00661  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00670  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: US 60/234,687  
;; PRIOR FILING DATE: 2000-09-21  
;; PRIOR APPLICATION NUMBER: US 09/608,408  
;; PRIOR FILING DATE: 2000-06-30  
;; PRIOR APPLICATION NUMBER: US 09/774,203  
;; PRIOR FILING DATE: 2001-01-29  
;; NUMBER OF SEQ ID NOS: 49117  
;; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
;; SEQ ID NO 48608  
;; LENGTH: 82  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens

FEATURE:  
OTHER INFORMATION: MAP TO AL118518.13  
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 7.6  
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 6.5  
OTHER INFORMATION: SWISSPROT HIT: P53693, EVALUATE 3.60e+00  
OTHER INFORMATION: EST\_HUMAN HIT: AW205702.1, EVALUATE 2.60e-01  
US-09-864-761-48608

Query Match 76.5%; Score 13; DB 9; Length 82;  
Best Local Similarity 50.0%; Pred. No. 1.8e+03;  
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7  
DB 21 ETSSLL 26

RESULT 13  
US-09-864-761-48010  
Sequence 48010, Application US/09864761  
Patent No. US20020048763A1  
GENERAL INFORMATION:  
APPLICANT: Penn, Sharon G.  
APPLICANT: Rank, David R.  
APPLICANT: Hanzel, David K.  
APPLICANT: Chen, Wensheng  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
FILE REFERENCE: Acomica-X-1  
CURRENT APPLICATION NUMBER: US/09/864,761  
PRIOR FILING DATE: 2001-05-23  
PRIOR APPLICATION NUMBER: US 60/180,312  
PRIOR FILING DATE: 2000-02-04  
PRIOR APPLICATION NUMBER: US 60/207,456  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: US 09/632,366  
PRIOR FILING DATE: 2000-08-03  
PRIOR APPLICATION NUMBER: GB 24263.6  
PRIOR FILING DATE: 2000-10-04  
PRIOR APPLICATION NUMBER: US 60/236,359  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: PCT/US01/00666  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00667  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00664  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00669  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00665  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00668  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00663  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00662  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00661  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00670  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: US 60/234,687  
PRIOR FILING DATE: 2000-09-21  
PRIOR APPLICATION NUMBER: US 09/608,408  
PRIOR FILING DATE: 2000-06-30  
PRIOR APPLICATION NUMBER: US 09/774,203  
PRIOR FILING DATE: 2001-01-29  
NUMBER OF SEQ ID NOS: 49117  
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
SEQ ID NO 48010  
LENGTH: 88  
TYPE: PRT  
ORGANISM: Homo sapiens

FEATURE:  
OTHER INFORMATION: MAP TO AC007038.3  
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.98  
OTHER INFORMATION: EST\_HUMAN HIT: T23219.1, EVALUATE 2.00e-35  
OTHER INFORMATION: SWISSPROT HIT: P17801, EVALUATE 1.10e+00  
US-09-864-761-48010

Query Match 76.5%; Score 13; DB 9; Length 88;  
Best Local Similarity 50.0%; Pred. No. 1.9e+03;  
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7  
DB 40 ETSSLL 45

RESULT 14  
US-10-106-698-5724  
Sequence 5724, Application US/10106698  
Publication No. US20030109690A1  
GENERAL INFORMATION:  
APPLICANT: Ruben et al.  
TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide  
FILE REFERENCE: PA005P1  
CURRENT APPLICATION NUMBER: US/10/106,698  
CURRENT FILING DATE: 2002-03-27  
PRIOR APPLICATION NUMBER: PCT/US00/26524  
PRIOR FILING DATE: 2000-09-28  
PRIOR APPLICATION NUMBER: US 60/157,137  
PRIOR FILING DATE: 1999-09-29  
PRIOR APPLICATION NUMBER: US 60/163,280  
PRIOR FILING DATE: 1999-11-03  
NUMBER OF SEQ ID NOS: 8564  
SOFTWARE: PatentIn Ver. 3.0  
SEQ ID NO 5724  
LENGTH: 94  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-106-698-5724

Query Match 76.5%; Score 13; DB 15; Length 94;  
Best Local Similarity 50.0%; Pred. No. 2e+03;  
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7  
DB 65 ETALL 70

RESULT 15  
US-09-864-761-37529  
Sequence 37529, Application US/09864761  
Patent No. US20020048763A1  
GENERAL INFORMATION:  
APPLICANT: Penn, Sharon G.  
APPLICANT: Rank, David R.  
APPLICANT: Hanzel, David K.  
APPLICANT: Chen, Wensheng  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
FILE REFERENCE: Acomica-X-1  
CURRENT APPLICATION NUMBER: US/09/864,761  
CURRENT FILING DATE: 2001-05-23  
PRIOR APPLICATION NUMBER: US 60/180,312  
PRIOR FILING DATE: 2000-02-04  
PRIOR APPLICATION NUMBER: US 60/207,456  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: US 09/632,366  
PRIOR FILING DATE: 2000-08-03  
PRIOR APPLICATION NUMBER: GB 24263.6  
PRIOR FILING DATE: 2000-10-04  
PRIOR APPLICATION NUMBER: US 60/236,359  
PRIOR FILING DATE: 2000-09-27



; PRIOR APPLICATION NUMBER: PCT/US01/00666  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00667  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00664  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00669  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00665  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00668  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00663  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00662  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00661  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00670  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: US 60/234,687  
 ; PRIOR FILING DATE: 2000-09-21  
 ; PRIOR APPLICATION NUMBER: US 09/608,408  
 ; PRIOR FILING DATE: 2000-06-30  
 ; PRIOR APPLICATION NUMBER: US 09/774,203  
 ; PRIOR FILING DATE: 2001-01-29  
 ; NUMBER OF SEQ ID NOS: 49117  
 ; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1  
 ; SEQ ID NO 37529  
 ; LENGTH: 98  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; OTHER INFORMATION: MAP TO AF003627.1  
 ; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 0.91  
 ; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.2  
 ; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 0.97  
 ; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.2  
 ; OTHER INFORMATION: SWISSPROT HIT: P34675, EVALUE 8.40e+00  
 ; OTHER INFORMATION: EST\_HUMAN HIT: AU143159.1, EVALUE 4.20e+00  
 US-09-864-761-37529

Query Match 76.5%; Score 13; DB 9; Length 98;  
 Best Local Similarity 50.0%; Pred. No. 2.1e+03;  
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7  
 60 EATSL 65

RESULT 16  
 US-09-738-626-3698  
 ; Sequence 3698, Application US/09738626  
 ; Publication No. US20020197605A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: NAKAGAWA, SATOSHI  
 ; APPLICANT: MIZOGUCHI, HIROSHI  
 ; APPLICANT: ANDO, SEIKO  
 ; APPLICANT: HAYASHI, MIKIRO  
 ; APPLICANT: OCHIAI, KEIKO  
 ; APPLICANT: YOKOI, HARUHIKO  
 ; APPLICANT: TATEISHI, NAKO  
 ; APPLICANT: SENOH, AKIHIRO  
 ; APPLICANT: IKEDA, MASATO  
 ; APPLICANT: OZAKI, AKIO  
 ; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
 ; FILE REFERENCE: 249-125  
 ; CURRENT APPLICATION NUMBER: US/09/738,626  
 ; CURRENT FILING DATE: 2000-12-18  
 ; PRIOR APPLICATION NUMBER: JP 99/377484  
 ; PRIOR FILING DATE: 1999-12-16  
 ; PRIOR APPLICATION NUMBER: JP 00/159162

; PRIOR FILING DATE: 2000-04-07  
 ; PRIOR APPLICATION NUMBER: JP 00/280988  
 ; PRIOR FILING DATE: 2000-08-03  
 ; NUMBER OF SEQ ID NOS: 7059  
 ; SOFTWARE: PatentIn ver. 3.0  
 ; SEQ ID NO 3698  
 ; LENGTH: 138  
 ; TYPE: PRT  
 ; ORGANISM: Corynebacterium glutamicum  
 US-09-738-626-3698

Query Match 76.5%; Score 13; DB 10; Length 138;  
 Best Local Similarity 50.0%; Pred. No. 2.9e+03;  
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7  
 36 ESAALL 41

RESULT 17  
 US-10-138-473-30  
 ; Sequence 30, Application US/10138473  
 ; Publication No. US20030165525A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: ANDERSEN, Peter  
 ; APPLICANT: WELTING, Karin  
 ; APPLICANT: HANSEN, Christina Veggerby  
 ; APPLICANT: FLORIO, Walter  
 ; APPLICANT: ORGELS, Li Mei Meng  
 ; APPLICANT: SKORT, Rikke Louise Vinther  
 ; APPLICANT: ROSENKRANDS, Ida  
 ; APPLICANT: RASMUSSEN, Peter Birk  
 ; TITLE OF INVENTION: TB Diagnostic Based On Antigens From M. tuberculosis  
 ; FILE REFERENCE: 0459-0710P  
 ; CURRENT APPLICATION NUMBER: US/10/138,473  
 ; CURRENT FILING DATE: 2002-05-02  
 ; PRIOR APPLICATION NUMBER: US 10/060,428  
 ; PRIOR FILING DATE: 2002-01-29  
 ; PRIOR APPLICATION NUMBER: US 09/415,884  
 ; PRIOR FILING DATE: 1999-10-08  
 ; PRIOR APPLICATION NUMBER: US 60/116,673  
 ; PRIOR FILING DATE: 1999-01-21  
 ; PRIOR APPLICATION NUMBER: DK 1998 01281  
 ; PRIOR FILING DATE: 1998-10-18  
 ; PRIOR APPLICATION NUMBER: US 60/070,488  
 ; PRIOR FILING DATE: 1998-01-05  
 ; PRIOR APPLICATION NUMBER: DK 1997 01277  
 ; PRIOR FILING DATE: 1997-11-10  
 ; PRIOR APPLICATION NUMBER: US 60/044,624  
 ; PRIOR FILING DATE: 1997-04-18  
 ; PRIOR APPLICATION NUMBER: DK 1997 00376  
 ; PRIOR FILING DATE: 1997-04-02  
 ; NUMBER OF SEQ ID NOS: 174  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 30  
 ; LENGTH: 144  
 ; TYPE: PRT  
 ; ORGANISM: Mycobacterium tuberculosis  
 US-10-138-473-30

Query Match 76.5%; Score 13; DB 12; Length 144;  
 Best Local Similarity 50.0%; Pred. No. 3e+03;  
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7  
 83 ESSLL 88

RESULT 18  
 US-10-080-170-94  
 ; Sequence 94, Application US/10080170

```

; Publication No. US20030129601A1
; GENERAL INFORMATION:
; APPLICANT: COLE, S.T.
; TITLE OF INVENTION: COMPARATIVE MYCOBACTERIAL GENOMICS AS A TOOL FOR
; IDENTIFYING TARGETS FOR THE DIAGNOSIS, PROPHYLAXIS OR
; TREATMENT OF MYCOBACTERIOSES
; FILE REFERENCE: 03495.0218
; CURRENT APPLICATION NUMBER: US/10/080,170
; PRIOR FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: 60/270,123
; NUMBER OF SEQ ID NOS: 652
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 94
; LENGTH: 144
; TYPE: PRT
; ORGANISM: Mycobacterium leprae
; US-10-080-170-94

Query Match      76.5%; Score 13; DB 16; Length 144;
Best Local Similarity 50.0%; Pred. No. 3e+03;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      2 EXXXLL 7
DB      83 ESSSL 88

RESULT 19
US-10-080-170-518
; Sequence 518, Application US/10080170
; Publication No. US20030129601A1
; GENERAL INFORMATION:
; APPLICANT: COLE, S.T.
; TITLE OF INVENTION: COMPARATIVE MYCOBACTERIAL GENOMICS AS A TOOL FOR
; IDENTIFYING TARGETS FOR THE DIAGNOSIS, PROPHYLAXIS OR
; TREATMENT OF MYCOBACTERIOSES
; FILE REFERENCE: 03495.0218
; CURRENT APPLICATION NUMBER: US/10/080,170
; PRIOR FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: 60/270,123
; NUMBER OF SEQ ID NOS: 652
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 518
; LENGTH: 144
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
; US-10-080-170-518

Query Match      76.5%; Score 13; DB 16; Length 144;
Best Local Similarity 50.0%; Pred. No. 3e+03;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      2 EXXXLL 7
DB      83 ESSSL 88

RESULT 20
US-09-739-254-67
; Sequence 67, Application US/09739254
; Patent No. US20010021700A1
; GENERAL INFORMATION:
; APPLICANT: ROSEN ET AL.
; TITLE OF INVENTION: 49 Human Secreted Proteins
; FILE REFERENCE: P2032P1
; CURRENT APPLICATION NUMBER: US/09/739,254
; CURRENT FILING DATE: 2000-12-19
; EARLIER APPLICATION NUMBER: 09/511,554
; EARLIER FILING DATE: 2000-02-23
; EARLIER APPLICATION NUMBER: PCT/US99/19330
; EARLIER FILING DATE: 1999-08-24

```

```

; EARLIER APPLICATION NUMBER: 60/097,917
; EARLIER FILING DATE: 1998-08-25
; EARLIER APPLICATION NUMBER: 60/098,634
; EARLIER FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 170
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 67
; LENGTH: 156
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (156)
; OTHER INFORMATION: Xaa equals stop translation
; US-09-739-254-67

```

```

Query Match      76.5%; Score 13; DB 9; Length 156;
Best Local Similarity 50.0%; Pred. No. 3.2e+03;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

QY      2 EXXXLL 7
DB      100 EAAALL 105

```

```

RESULT 21
US-09-904-615-67
; Sequence 67, Application US/09904615
; Patent No. US20020026040A1
; GENERAL INFORMATION:
; APPLICANT: ROSEN ET AL.
; TITLE OF INVENTION: 49 Human Secreted Proteins
; FILE REFERENCE: P2032P1
; CURRENT APPLICATION NUMBER: US/09/904,615
; CURRENT FILING DATE: 2001-07-16
; PRIOR APPLICATION NUMBER: 09/511,554
; PRIOR FILING DATE: 2000-02-23
; PRIOR APPLICATION NUMBER: 60/097,917
; PRIOR FILING DATE: 1998-08-25
; PRIOR APPLICATION NUMBER: 60/098,634
; PRIOR FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 170
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 67
; LENGTH: 156
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (156)
; OTHER INFORMATION: Xaa equals stop translation
; US-09-904-615-67

```

```

Query Match      76.5%; Score 13; DB 9; Length 156;
Best Local Similarity 50.0%; Pred. No. 3.2e+03;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

QY      2 EXXXLL 7
DB      100 EAAALL 105

```

```

RESULT 22
US-10-055-098-67
; Sequence 67, Application US/10055098
; Publication No. US20030139954A1
; GENERAL INFORMATION:
; APPLICANT: ROSEN ET AL.
; TITLE OF INVENTION: 49 Human Secreted Proteins
; FILE REFERENCE: P2032P1
; CURRENT APPLICATION NUMBER: US/10/055,098
; CURRENT FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/511,554

```

```

; PRIOR FILING DATE: EARLIER FILING DATE: 2000-02-23
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: PCT/US99/19330
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-08-24
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/097,917
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-08-25
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/098,634
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 170
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 67
; LENGTH: 156
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (156)
; OTHER INFORMATION: Xaa equals stop translation
US-10-054-988-67

Query Match          76.5%; Score 13; DB 12; Length 156;
Best Local Similarity 50.0%; Pred. No. 3.2e+03;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7
DB 100 EAAALL 105

RESULT 23
US-10-054-988-67
; Sequence 67, Application US/10054988
; Publication No. US20030087341A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 49 Human Secreted Proteins
; FILE REFERENCE: P2032P1
; CURRENT APPLICATION NUMBER: US/10/054,988
; CURRENT FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: 09/904,615
; PRIOR FILING DATE: 2001-07-16
; PRIOR APPLICATION NUMBER: 09/511,554
; PRIOR FILING DATE: 2000-02-23
; PRIOR APPLICATION NUMBER: 60/097,917
; PRIOR FILING DATE: 1998-08-25
; PRIOR APPLICATION NUMBER: 60/098,634
; PRIOR FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 170
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 67
; LENGTH: 156
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (156)
; OTHER INFORMATION: Xaa equals stop translation
US-10-054-988-67

Query Match          76.5%; Score 13; DB 15; Length 156;
Best Local Similarity 50.0%; Pred. No. 3.2e+03;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7
DB 100 EAAALL 105

RESULT 24
US-09-840-787-15
; Sequence 15, Application US/09840787
; Patent No. US20020058264A1
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC006
; CURRENT APPLICATION NUMBER: US/09/764,891
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 10231
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4552
; LENGTH: 205
; TYPE: PRT
```

```

; Hillman, Jennifer L.
; Bandman, Olga
; Shah, Purvi
; Au-Young, Janice
; Yue, Henry
; Guegler, Karl J.
; Corley, Neil C.
; TITLE OF INVENTION: HUMAN REGULATORY MOLECULES
; NUMBER OF SEQUENCES: 98
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/840,787
; FILING DATE: 23-Apr-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/518,865
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0356 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 158 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: BRSTTUT01
; CLONE: 606742
; SEQUENCE DESCRIPTION: SEQ ID NO: 15 :
US-09-840-787-15

Query Match          76.5%; Score 13; DB 9; Length 158;
Best Local Similarity 50.0%; Pred. No. 3.3e+03;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7
DB 74 EAAALL 79

RESULT 25
US-09-764-891-4552
; Sequence 4552, Application US/09764891
; Publication No. US20030077808A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC006
; CURRENT APPLICATION NUMBER: US/09/764,891
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 10231
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4552
; LENGTH: 205
; TYPE: PRT
```

```

; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (17)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-891-4552

Query Match          76.5%; Score 13; DB 11; Length 205;
Best Local Similarity 50.0%; Pred. No. 4.2e+03;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7
DB 24 EAAALL 29

RESULT 26
US-10-046-938-2
Sequence 2, Application US/10046938
Publication No. US20020192185A1
GENERAL INFORMATION:
APPLICANT: MITTAL, SURESH K.
APPLICANT: GRAHAM, FRANK L.
APPLICANT: PREVIC, LUDVIG
APPLICANT: BABIUK, LORNE A.
TITLE OF INVENTION: MAMMALIAN CELL LINES EXPRESSING BOVINE ADENOVIRUS FUNCTIONS
FILE REFERENCE: 293102002102
CURRENT APPLICATION NUMBER: US/10/046,938
CURRENT FILING DATE: 2002-01-14
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/435,242
PRIOR FILING DATE: EARLIER FILING DATE: 1999-11-05
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 08/815,927
PRIOR FILING DATE: EARLIER FILING DATE: 1997-03-13
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 08/164,294
PRIOR FILING DATE: EARLIER FILING DATE: 1993-12-09
NUMBER OF SEQ ID NOS: 34
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 211
TYPE: PRT
ORGANISM: Bovine adenovirus type 3
US-10-046-938-2

Query Match          76.5%; Score 13; DB 14; Length 211;
Best Local Similarity 50.0%; Pred. No. 4.3e+03;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

DB 2 EXXXLL 7
70 EASSLL 75

RESULT 27
US-10-029-386-32343
Sequence 32343, Application US/10029386
Publication No. US20030194704A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharron G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR C
TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
FILE REFERENCE: AEMICA-X-2
CURRENT APPLICATION NUMBER: US/10/029,386
CURRENT FILING DATE: 2001-12-20
NUMBER OF SEQ ID NOS: 34288
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
SEQ ID NO 32343
LENGTH: 224
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AL031387.4

```

```

; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.7
; OTHER INFORMATION: SWISSPROT HIT: P51841, EVALUATE 1.00e-130
US-10-029-386-32343

```

```

Query Match          76.5%; Score 13; DB 12; Length 224;
Best Local Similarity 50.0%; Pred. No. 4.5e+03;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

QY 2 EXXXLL 7
DB 114 EAAALL 119

```

```

RESULT 28
US-09-893-737-184
Sequence 184, Application US/09893737
Patent No. US20020110855A1
GENERAL INFORMATION:
APPLICANT: Sheppard, Paul O.
APPLICANT: Presnell, Scott R.
TITLE OF INVENTION: MAMMALIAN SECRETED PROTEINS
FILE REFERENCE: 00-41
CURRENT APPLICATION NUMBER: US/09/893,737
CURRENT FILING DATE: 2001-06-28
PRIOR APPLICATION NUMBER: US 60/215,446
PRIOR FILING DATE: 2000-06-30
NUMBER OF SEQ ID NOS: 329
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 184
LENGTH: 230
TYPE: PRT
ORGANISM: Homo sapiens
US-09-893-737-184

```

```

Query Match          76.5%; Score 13; DB 10; Length 230;
Best Local Similarity 50.0%; Pred. No. 4.7e+03;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

QY 2 EXXXLL 7
DB 183 EATSLL 188

```

```

RESULT 29
US-09-764-891-2999
Sequence 2999, Application US/09764891
Publication No. US20030077808A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PC006
CURRENT APPLICATION NUMBER: US/09/764,891
CURRENT FILING DATE: 2001-01-17
Prior application data removed - consult PALM or file wrapper
NUMBER OF SEQ ID NOS: 10231
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2999
LENGTH: 234
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (199)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (205)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

```

; NAME/KEY: SITE  
; LOCATION: (230)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-09-764-891-2999

Query Match 76.5%; Score 13; DB 11; Length 234;  
Best Local Similarity 50.0%; Pred. No. 4.7e+03;  
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 EXXXLL 7  
| | |  
Db 215 EASLL 220

RESULT 30  
US-09-861-289-18  
; Sequence 18, Application US/09861289  
; Patent No. US20020110897A1  
; GENERAL INFORMATION:  
; APPLICANT: Sherman, D.H.  
; APPLICANT: Liu, H.  
; APPLICANT: Xue, Y.  
; APPLICANT: Zhao, L.  
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin  
; FILE REFERENCE: 600.438US1  
; CURRENT APPLICATION NUMBER: US/09/861,289  
; CURRENT FILING DATE: 2001-05-18  
; PRIOR APPLICATION NUMBER: 09/105,537  
; PRIOR FILING DATE: 1998-06-26  
; NUMBER OF SEQ ID NOS: 43  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 18  
; LENGTH: 237  
; TYPE: PRT  
; ORGANISM: Streptomyces venezuelae  
US-09-861-289-18

Query Match 76.5%; Score 13; DB 10; Length 237;  
Best Local Similarity 50.0%; Pred. No. 4.8e+03;  
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 EXXXLL 7  
| | |  
Db 38 EASLL 43

RESULT 31  
US-09-860-846-18  
; Sequence 18, Application US/09860846  
; Patent No. US20020164742A1  
; GENERAL INFORMATION:  
; APPLICANT: Sherman, D.H.  
; APPLICANT: Liu, H.  
; APPLICANT: Xue, Y.  
; APPLICANT: Zhao, L.  
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin  
; FILE REFERENCE: 600.438US1  
; CURRENT APPLICATION NUMBER: US/09/860,846  
; CURRENT FILING DATE: 2001-05-18  
; PRIOR APPLICATION NUMBER: 09/105,537  
; PRIOR FILING DATE: 1998-06-26  
; NUMBER OF SEQ ID NOS: 43  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 18  
; LENGTH: 237  
; TYPE: PRT  
; ORGANISM: Streptomyces venezuelae  
US-09-860-846-18

Query Match 76.5%; Score 13; DB 10; Length 237;  
Best Local Similarity 50.0%; Pred. No. 4.8e+03;  
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 EXXXLL 7  
| | |  
Db 38 EASLL 43

RESULT 32  
US-09-988-384B-18  
; Sequence 18, Application US/09988384B  
; Publication No. US20030073824A1  
; GENERAL INFORMATION:  
; APPLICANT: Sherman, D.H.  
; APPLICANT: Liu, H.  
; APPLICANT: Xue, Y.  
; APPLICANT: Zhao, L.  
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin  
; FILE REFERENCE: 600.536US1  
; CURRENT APPLICATION NUMBER: US/09/988,384B  
; CURRENT FILING DATE: 2001-11-19  
; PRIOR APPLICATION NUMBER: PCT/US99/14398  
; PRIOR FILING DATE: 1999-06-25  
; PRIOR APPLICATION NUMBER: US 09/105,537  
; PRIOR FILING DATE: 1998-06-26  
; NUMBER OF SEQ ID NOS: 53  
; SEQ ID NO 18  
; LENGTH: 237  
; TYPE: PRT  
; ORGANISM: Streptomyces venezuelae  
US-09-988-384B-18

Query Match 76.5%; Score 13; DB 11; Length 237;  
Best Local Similarity 50.0%; Pred. No. 4.8e+03;  
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 EXXXLL 7  
| | |  
Db 38 EASLL 43

RESULT 33  
US-09-836-821-18  
; Sequence 18, Application US/09836821  
; Publication No. US20030087405A1  
; GENERAL INFORMATION:  
; APPLICANT: Sherman, D.H.  
; APPLICANT: Liu, H.  
; APPLICANT: Xue, Y.  
; APPLICANT: Zhao, L.  
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin  
; FILE REFERENCE: 600.438US1  
; CURRENT APPLICATION NUMBER: US/09/836,821  
; CURRENT FILING DATE: 2001-04-17  
; PRIOR APPLICATION NUMBER: 09/105,537  
; PRIOR FILING DATE: 1998-06-26  
; NUMBER OF SEQ ID NOS: 43  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 18  
; LENGTH: 237  
; TYPE: PRT  
; ORGANISM: Streptomyces venezuelae  
US-09-836-821-18

Query Match 76.5%; Score 13; DB 11; Length 237;  
Best Local Similarity 50.0%; Pred. No. 4.8e+03;  
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 EXXXLL 7  
| | |  
Db 38 EASLL 43

RESULT 34  
US-09-793-708-11  
; Sequence 11, Application US/09793708

Publication No. US20030104597A1  
GENERAL INFORMATION:  
APPLICANT: ASHLEY, Gary  
APPLICANT: BETLACH, Melanie C.  
APPLICANT: BETLACH, Mary C.  
APPLICANT: MCDANIEL, Robert  
APPLICANT: TANG, Li  
TITLE OF INVENTION: RECOMBINANT NARBONOLIDE POLYKETIDE SYNTHASE  
FILE REFERENCE: 30622002121  
CURRENT APPLICATION NUMBER: US/09/793,708  
CURRENT FILING DATE: 2001-02-22  
PRIOR APPLICATION NUMBER: US/09/657,440  
PRIOR FILING DATE: 2000-09-07  
PRIOR APPLICATION NUMBER: US/09/320,878  
PRIOR FILING DATE: 1999-05-27  
PRIOR APPLICATION NUMBER: US/09/141,908  
PRIOR FILING DATE: 1998-08-28  
PRIOR APPLICATION NUMBER: US/09/073,538  
PRIOR FILING DATE: 1998-05-06  
PRIOR APPLICATION NUMBER: US/08/846,247  
PRIOR FILING DATE: 1997-04-30  
PRIOR APPLICATION NUMBER: US/60/134,990  
PRIOR FILING DATE: 1999-05-20  
NUMBER OF SEQ ID NOS: 38  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 11  
LENGTH: 237  
TYPE: PRT  
ORGANISM: Streptomyces venezuelae  
US-09-793-708-11

Query Match 76.5%; Score 13; DB 11; Length 237;  
Best Local Similarity 50.0%; Pred. No. 4.8e+03;  
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7  
DB 38 EASSLL 43

RESULT 35  
US-10-160-539-11  
Sequence 11, Application US/10160539  
Publication No. US20030162262A1  
GENERAL INFORMATION:  
APPLICANT: ASHLEY, Gary  
APPLICANT: BETLACH, Melanie C.  
APPLICANT: BETLACH, Mary C.  
APPLICANT: MCDANIEL, Robert  
APPLICANT: TANG, Li  
TITLE OF INVENTION: RECOMBINANT NARBONOLIDE POLYKETIDE SYNTHASE  
FILE REFERENCE: 30622002120  
CURRENT APPLICATION NUMBER: US/10/160,539  
CURRENT FILING DATE: 2002-05-29  
PRIOR APPLICATION NUMBER: US/09/657,440  
PRIOR FILING DATE: 2000-09-07  
PRIOR APPLICATION NUMBER: US/09/320,878  
PRIOR FILING DATE: 1999-05-27  
PRIOR APPLICATION NUMBER: CIP OF 09/141,908  
PRIOR FILING DATE: 1998-08-28  
NUMBER OF SEQ ID NOS: 34  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 11  
LENGTH: 237  
TYPE: PRT  
ORGANISM: Streptomyces venezuelae  
US-10-160-539-11

Query Match 76.5%; Score 13; DB 12; Length 237;  
Best Local Similarity 50.0%; Pred. No. 4.8e+03;  
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7

Db 38 EASSLL 43

RESULT 36  
US-10-271-889-18  
Sequence 18, Application US/10271889  
Publication No. US20030194784A1  
GENERAL INFORMATION:  
APPLICANT: Sherman, D.H.  
APPLICANT: Liu, H.  
APPLICANT: Xue, Y.  
APPLICANT: Zhao, L.  
TITLE OF INVENTION: DNA Encoding Methymycin and Pikromycin  
FILE REFERENCE: 600.582US1  
CURRENT APPLICATION NUMBER: US/10/271,889  
CURRENT FILING DATE: 2002-10-15  
PRIOR APPLICATION NUMBER: US/09/861,289  
PRIOR FILING DATE: 2001-05-18  
PRIOR APPLICATION NUMBER: US/09/860,846  
PRIOR FILING DATE: 2001-05-18  
PRIOR APPLICATION NUMBER: US/09/836,821  
PRIOR FILING DATE: 2001-04-17  
PRIOR APPLICATION NUMBER: US/09/105,537  
PRIOR FILING DATE: 1998-06-26  
NUMBER OF SEQ ID NOS: 55  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 18  
LENGTH: 237  
TYPE: PRT  
ORGANISM: Streptomyces venezuelae  
US-10-271-889-18

Query Match 76.5%; Score 13; DB 12; Length 237;  
Best Local Similarity 50.0%; Pred. No. 4.8e+03;  
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7  
DB 38 EASSLL 43

RESULT 37  
US-09-858-664A-7  
Sequence 7, Application US/09858664A  
Patent No. US20020072491A1  
GENERAL INFORMATION:  
APPLICANT: WEI, Ming-Hui, et al.  
TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC  
ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES  
THEREOF  
FILE REFERENCE: CLO00927-CIP  
CURRENT APPLICATION NUMBER: US/09/858,664A  
CURRENT FILING DATE: 2001-05-17  
PRIOR APPLICATION NUMBER: 09/711,134  
PRIOR FILING DATE: 2000-11-11  
NUMBER OF SEQ ID NOS: 33  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 7  
LENGTH: 245  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-858-664A-7

Query Match 76.5%; Score 13; DB 9; Length 245;  
Best Local Similarity 50.0%; Pred. No. 4.9e+03;  
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7  
DB 48 EAAATLL 53

## RESULT 38

US-09-815-242-13461  
; Sequence 13461, Application US/09815242  
; Patent No. US20020061569A1  
; GENERAL INFORMATION:  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari L.  
; APPLICANT: Zyskind, Judith W.  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John D.  
; APPLICANT: Carr, Grant J.  
; APPLICANT: Yamamoto, Robert T.  
; APPLICANT: Xu, H. Howard  
; TITLE OF INVENTION: Identification of Essential Genes in  
; FILE REFERENCE: ELITRA.011A  
; CURRENT APPLICATION NUMBER: US/09/815,242  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; NUMBER OF SEQ ID NOS: 14110  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 13461  
; LENGTH: 247  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
US-09-815-242-13461

Query Match 76.5%; Score 13; DB 9; Length 247;  
Best Local Similarity 50.0%; Pred. No. 5e+03;  
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7  
| | |  
Db 176 EAAALL 181

## RESULT 39

US-09-815-242-13622  
; Sequence 13622, Application US/09815242  
; Patent No. US20020061569A1  
; GENERAL INFORMATION:  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari L.  
; APPLICANT: Zyskind, Judith W.  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John D.  
; APPLICANT: Carr, Grant J.  
; APPLICANT: Yamamoto, Robert T.  
; APPLICANT: Xu, H. Howard  
; TITLE OF INVENTION: Identification of Essential Genes in  
; FILE REFERENCE: ELITRA.011A  
; CURRENT APPLICATION NUMBER: US/09/815,242  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26

Query Match 76.5%; Score 13; DB 9; Length 247;  
Best Local Similarity 50.0%; Pred. No. 5e+03;  
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7  
| | |  
Db 176 EAAALL 181

; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; NUMBER OF SEQ ID NOS: 14110  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 13622  
; LENGTH: 247  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
US-09-815-242-13622

Query Match 76.5%; Score 13; DB 9; Length 247;  
Best Local Similarity 50.0%; Pred. No. 5e+03;  
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7  
| | |  
Db 176 EAAALL 181

## RESULT 40

US-10-024-632-13  
; Sequence 13, Application US/10024632  
; Publication No. US20020170093A1  
; GENERAL INFORMATION:  
; APPLICANT: Monsanto Technology LLC  
; APPLICANT: He, Steve S.  
; APPLICANT: Dotson, Stanton B.  
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ASSOCIATED WITH PLANT CELL PROLIFERATION  
; FILE REFERENCE: 38-21(51837)B  
; CURRENT APPLICATION NUMBER: US/10/024,632  
; PRIOR FILING DATE: 2001-12-19  
; PRIOR APPLICATION NUMBER: US 60/257,896  
; PRIOR FILING DATE: 2000-12-21  
; NUMBER OF SEQ ID NOS: 33  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 13  
; LENGTH: 255  
; TYPE: PRT  
; ORGANISM: Zea mays  
US-10-024-632-13

Query Match 76.5%; Score 13; DB 14; Length 255;  
Best Local Similarity 50.0%; Pred. No. 5.1e+03;  
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7  
| | |  
Db 88 ESSTLL 93

Search completed: November 5, 2003, 14:23:15  
Job time : 29 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

# OM protein - protein search, using sw model

Run on: November 5, 2003, 14:15:08, Search time 21 Seconds  
(without alignments)  
32.056 Million cell updates/sec

Title: US-09-620-840C-2

Perfect score: 17

Sequence: 1 EXXXLL 7

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Seatched: 293308 seqs, 96168682 residues

1 number of hits satisfying chosen parameters: 293308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

## Database :

PIR 76:\*

1: piri:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	13	76.5	91	2 H96701	hypothetical prote
2	13	76.5	98	2 B75508	hypothetical prote
3	13	76.5	107	2 C81172	hypothetical prote
4	13	76.5	107	2 B81935	hypothetical prote
5	13	76.5	109	1 PVFGA	parvalbumin alpha
6	13	76.5	110	2 S11054	hypothetical prote
7	13	76.5	114	2 F82782	hypothetical prote
8	13	76.5	115	2 S69533	hypothetical prote
9	13	76.5	118	2 A71240	hypothetical prote
10	13	76.5	121	2 E81052	hypothetical prote
11	13	76.5	121	2 A72776	hypothetical prote
12	13	76.5	123	2 S73426	hypothetical prote
13	13	76.5	129	2 B72669	probable lipoprote
14	13	76.5	134	1 S63331	hypothetical prote
15	13	76.5	136	2 E36790	probable membrane
16	13	76.5	136	2 T48326	hypothetical prote
17	13	76.5	138	2 S70820	hypothetical prote
18	13	76.5	140	2 H90899	H+-transporting tw
19	13	76.5	140	2 G90997	minor tail protein
20	13	76.5	140	2 F85817	probable minor tail
21	13	76.5	141	2 H64361	probable minor tail
22	13	76.5	144	2 C70937	hypothetical prote
23	13	76.5	144	2 C87020	hypothetical prote
24	13	76.5	144	2 A81824	conserved hypotet
25	13	76.5	151	2 D70007	hypothetical prote
26	13	76.5	152	2 E72487	hypothetical prote
27	13	76.5	161	2 C75438	hypothetical prote
28	13	76.5	163	2 G70827	hypothetical prote
29	13	76.5	166	2 T36616	hypothetical prote

30	13	76.5	176	2 G86796	prophase p13 prote
31	13	76.5	177	2 S54778	NR-13 protein - qu
32	13	76.5	177	2 AP0438	conserved hypotet
33	13	76.5	179	2 C70507	hypothetical prote
34	13	76.5	180	2 G82858	hypothetical prote
35	13	76.5	181	2 B87138	probable adenylate
36	13	76.5	183	2 S52661	ubiquitin-protein
37	13	76.5	184	2 B72383	ubiquitin-protein
38	13	76.5	184	2 T50342	thymidine kinase (
39	13	76.5	185	2 S43785	ubiquitin conjugat
40	13	76.5	186	2 D97709	ubiquitin-protein
41	13	76.5	187	2 B69314	hypothetical prote
42	13	76.5	190	2 AH2529	hypothetical prote
43	13	76.5	194	2 JN0645	transcription regu
44	13	76.5	194	2 D55217	naphthalene 1,2-di
45	13	76.5	194	2 JS0072	polycyclic aromati
					naphthalene dioxyg

## ALIGNMENTS

### RESULT 1

H96701  
hypothetical protein T23K23.24 [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001  
C:Accession: H96701  
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, C.; Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Dewar, K.; Huizar, L.; Jensen, R.F.; Hughes, B.; Huizar, L.; Nature 408, 816-820, 2000  
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C. C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A:Authors: Salberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A:Reference number: A86141; MUID:21016719; PMID:11130712  
A:Accession: H96701  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-91 <STO>  
A:Cross-references: GB:AE005173; NID:G6553904; PIDN:AAF16570.1; GSPDB:GN00141  
C:Genetics:  
A:Gene: T23K23.24  
A:Map position: 1

Query Match 76.5%; Score 13; DB 2; Length 91;  
Best Local Similarity 50.0%; Pred. No. 4.2e+02;  
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7  
DB 72 EAASLL 77

### RESULT 2

B75508  
hypothetical protein - Deinococcus radiodurans (strain R1)  
C:Species: Deinococcus radiodurans  
C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 31-Mar-2000  
C:Accession: B75508  
R:White, O.; Eisen, J.A.; Heidelberg, J.P.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.; M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma, S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.  
Science 286, 1571-1577, 1999

A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.  
A:Reference number: A75250; MUID:20036896; PMID:10567266  
A:Accession: B75508  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-98 <WHI>  
A:Cross-references: GB:AE001911; GB:AE000513; NID:G6458217; PIDN:AAF10113.1; PID:G645822



A;Experimental source: strain R1  
 C;Genetics:  
 A;Gene: DR0532  
 A;Map position: 1

Query Match 76.5%; Score 13; DB 2; Length 98;  
 Best Local Similarity 50.0%; Pred. No. 4.9e+02;  
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7  
 |  
 |  
 |  
 Db 43 ETTALL 48

## RESULT 3

hypothetical protein NMB0685 [imported] - Neisseria meningitidis (strain MC58 serogroup C)  
 C;Species: Neisseria meningitidis  
 C;Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 19-Jan-2001  
 C;Accession: C81172  
 A;Title: H. J. Saunders, N. J. Heidelberg, J. J. Jeffries, A. C. Nelson, K. B. Eisen, J. A. White, O. Fleischmann, R. D. Dougherty, B. A. J. H. Qin, H. Vamathevan, J. J. Gill, J. Scarlato, V. Masignani, V. Pizza, M. Science 287, 1809-1815, 2000  
 A;Authors: Grandi, G.; Sun, L.; Smith, H. O.; Fraser, C. M.; Moxon, E. R.; Rappuoli, R.; V. A;Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.  
 A;Reference number: A81000; MUID:20175755; PMID:10710307  
 A;Accession: C81172  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-107 <TST>  
 A;Cross-references: GB:AB002422; GB:AB002098; NID:G7225898; PIDN:AAF41103.1; PID:G7225898  
 A;Experimental source: serogroup B, strain MC58  
 C;Genetics:  
 A;Gene: NMB0685  
 C;Superfamily: Neisseria meningitidis hypothetical protein NMB0685

Query Match 76.5%; Score 13; DB 2; Length 107;  
 Best Local Similarity 50.0%; Pred. No. 4.9e+02;  
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7  
 |  
 |  
 |  
 Db 76 ETTALL 81

## RESULT 4

hypothetical protein NMA0887 [imported] - Neisseria meningitidis (strain Z2491 serogroup C)  
 C;Species: Neisseria meningitidis  
 C;Date: 05-May-2000 #sequence\_revision 05-May-2000 #text\_change 02-Feb-2001  
 C;Accession: B81935  
 A;Title: H. J. Saunders, N. J. Heidelberg, J. J. Jeffries, A. C. Nelson, K. B. Eisen, J. A. White, O. Fleischmann, R. D. Dougherty, B. A. J. H. Qin, H. Vamathevan, J. J. Gill, J. Scarlato, V. Masignani, V. Pizza, M. Nature 404, 502-506, 2000  
 A;Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.  
 A;Reference number: A81775; MUID:20222556; PMID:10761919  
 A;Accession: B81935  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-107 <PAR>  
 A;Cross-references: GB:AL162754; GB:AL157959; NID:G7379424; PIDN:CAB84167.1; PID:G737960  
 A;Experimental source: serogroup A, strain Z2491  
 C;Genetics:  
 A;Gene: NMA0887  
 C;Superfamily: Neisseria meningitidis hypothetical protein NMB0685

Query Match 76.5%; Score 13; DB 2; Length 107;  
 Best Local Similarity 50.0%; Pred. No. 4.9e+02;  
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7  
 |  
 |  
 |

Db 76 ETTALL 81

## RESULT 5

parvalbumin alpha - edible frog  
 C;Species: Rana esculenta (edible frog)  
 C;Date: 18-Aug-1992 #sequence\_revision 18-Aug-1992 #text\_change 01-Sep-1995  
 C;Accession: A03062  
 R;Jauregui-Adell, J.; Pechere, J. P.; Briand, G.; Richet, C.; Demaille, J. G. Eur. J. Biochem. 123, 337-345, 1992  
 A;Title: Amino-acid sequence of an alpha-parvalbumin, pI=4.88, from frog skeletal muscle  
 A;Reference number: A03062; MUID:82186701; PMID:7042341  
 A;Accession: A03062  
 A;Molecule type: protein  
 A;Residues: 1-109 <JAU>  
 C;Superfamily: parvalbumin; calmodulin repeat homology  
 C;Keywords: calcium binding; duplication; EF hand; muscle  
 F;38-70/Domain: calmodulin repeat homology <EF1>  
 F;77-109/Domain: calmodulin repeat homology <EF2>  
 F;51,53,55,57,59,62/Binding site: calcium (Asp, Asp, Ser, Phe, Glu, Glu) #status predicted  
 F;90,92,94,96,101/Binding site: calcium (Asp, Asp, Asp, Lys, Glu) #status predicted

Query Match 76.5%; Score 13; DB 1; Length 109;  
 Best Local Similarity 50.0%; Pred. No. 5e+02;  
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7  
 |  
 |  
 |  
 Db 81 ETTALL 86

## RESULT 6

parvalbumin - bullfrog  
 C;Species: Rana catesbeiana (bullfrog)  
 C;Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 18-Jun-1999  
 C;Accession: S11054  
 R;Sasaki, T.; Tanokura, M.; Asaoka, K. FEBS Lett. 268, 249-251, 1990  
 A;Title: The complete amino acid sequence of bullfrog (Rana catesbeiana) parvalbumin  
 A;Reference number: S11054; MUID:90346151; PMID:2384163  
 A;Accession: S11054  
 A;Status: preliminary  
 A;Molecule type: protein  
 A;Residues: 1-110 <SAS>  
 C;Superfamily: parvalbumin; calmodulin repeat homology  
 C;Keywords: EF hand  
 F;39-71/Domain: calmodulin repeat homology <EF1>  
 F;78-110/Domain: calmodulin repeat homology <EF2>

Query Match 76.5%; Score 13; DB 2; Length 110;  
 Best Local Similarity 50.0%; Pred. No. 5e+02;  
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7  
 |  
 |  
 |  
 Db 82 ETTALL 87

## RESULT 7

hypothetical protein XF0632 [imported] - Xylella fastidiosa (strain 9a5c)  
 C;Species: Xylella fastidiosa  
 C;Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 20-Aug-2000  
 C;Accession: F82782  
 R;Anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequence  
 A;Title: The genome sequence of the plant pathogen Xylella fastidiosa.  
 A;Reference number: A82515; MUID:20365717; PMID:10910347  
 A;Note: for a complete list of authors see reference number A59328 below  
 A;Accession: F82782  
 A;Status: preliminary

Query Match 76.5%; Score 13; DB 2; Length 107;  
 Best Local Similarity 50.0%; Pred. No. 4.9e+02;  
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7  
 |  
 |  
 |

A:Molecule type: DNA  
A:Residues: 1-114 <SIM>  
A:Cross-references: GB:AE003308; GB:AE003849; NID:G9105496; PIDN:AAF83442.1; GSPDB:GN001  
A:Experimental source: strain 945c  
R:Simpson, A.J.G.; Reinach, P.C.; Arruda, P.; Abreu, P.A.; Acencio, M.; Alvarenga, R.; A  
Briones, M.R.S.; Bueno, M.R.P.; Canargo, L.B.A.; Carraro, D.M.; Carrer, H  
as-Neto, E.; Docena, C.; El-Dorri, H.; Facinani, A.P.; Ferreira, A.J.S.  
submitted to GenBank, June 2000  
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm  
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laig  
Chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E  
A:Authors: Martins, E.M.F.; Matukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;  
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A  
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak  
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir  
M.; Tshukako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z  
A:Reference number: A59328  
A:Contents: annotation  
C:Genetics:  
ne: XF0632

Query Match 76.5%; Score 13; DB 2; Length 114;  
Best Local Similarity 50.0%; Pred. No. 5.2e+02;  
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 2 EXXXLL 7  
DB 49 EAAALL 54

RESULT 8  
S69533  
hypothetical protein 25 - phage HPI  
C:Species: phage HPI  
C:Date: 06-Dec-1996 #sequence\_revision 13-Mar-1997 #text\_change 20-Apr-2000  
C:Accession: S69533  
R:Reposito, D.; Fitzmaurice, W.P.; Benjamin, R.C.; Goodman, S.D.; Waldman, A.S.; Scocca,  
Nucleic Acids Res. 24, 2360-2368, 1996  
A:Title: The complete nucleotide sequence of bacteriophage HPI DNA.  
A:Reference number: S69503; MUID:96279738; PMID:8710508  
A:Accession: S69533  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-115 <ESP>  
A:Cross-references: EMBL:U24159; NID:G1046235; PIDN:AA09212.1; PID:G1046254  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, April 1995  
C:Superfamily: phage HPI hypothetical protein 25

Query Match 76.5%; Score 13; DB 2; Length 115;  
Best Local Similarity 50.0%; Pred. No. 5.2e+02;  
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 2 EXXXLL 7  
DB 35 ETTALL 40

RESULT 9  
A71240  
hypothetical protein PH0179 - Pyrococcus horikoshii  
C:Species: Pyrococcus horikoshii  
C:Date: 14-Aug-1998 #sequence\_revision 14-Aug-1998 #text\_change 18-Aug-2000  
C:Accession: A71240  
R:Kawarayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Sekin  
M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguchi  
DNA Res. 5, 55-76, 1998  
A:Title: Complete sequence and gene organization of the genome of a hyper-thermophilic a  
A:Reference number: A71000; MUID:98344137; PMID:9679194  
A:Accession: A71240  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-118 <KAW>  
A:Cross-references: GB:AP000001; NID:G3236128; PIDN:BAA29248.1; PID:G3256565

A:Experimental source: strain OT3  
A:Note: this accession replaces an interim accession for a sequence replaced by GenBank  
C:Genetics:  
A:Gene: PH0179  
C:Superfamily: Pyrococcus horikoshii hypothetical protein PH0179

Query Match 76.5%; Score 13; DB 2; Length 118;  
Best Local Similarity 50.0%; Pred. No. 5.4e+02;  
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 2 EXXXLL 7  
DB 46 ETSSLL 51

RESULT 10  
E81052  
hypothetical protein NMB1699 [imported] - Neisseria meningitidis (strain MCS8 serogroup  
C:Species: Neisseria meningitidis  
C:Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 19-Jan-2001  
C:Accession: E81052  
R:Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.J.  
Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;  
ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Masignani, V.; Pizza, M.  
Science 287, 1809-1815, 2000  
A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; V  
A:Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MCS8.  
A:Reference number: A81000; MUID:20175755; PMID:10710307  
A:Accession: E81052  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-121 <TET>  
A:Cross-references: GB:AE002520; GB:AE002520; NID:G7226947; PIDN:AAF42047.1; PID:G72269;  
A:Experimental source: serogroup B, strain MCS8  
C:Genetics:  
A:Gene: NMB1699  
C:Superfamily: Neisseria meningitidis hypothetical protein NMB1699

Query Match 76.5%; Score 13; DB 2; Length 121;  
Best Local Similarity 50.0%; Pred. No. 5.5e+02;  
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 2 EXXXLL 7  
DB 20 EASLL 25

RESULT 11  
A72776  
hypothetical protein APE0195 - Aeropyrum pernix (strain K1)  
C:Species: Aeropyrum pernix  
C:Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 09-Jun-2000  
C:Accession: A72776  
R:Kawarayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takat  
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; F  
DNA Res. 6, 83-101, 1999  
A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr  
A:Reference number: A72450; MUID:99310339; PMID:10382966  
A:Accession: A72776  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-121 <KAW>  
A:Cross-references: DDBJ:AP000058; NID:95103388; PIDN:BAA79107.1; PID:d1042883; PID:G51C  
A:Experimental source: strain K1  
C:Genetics:  
A:Gene: APE0195  
C:Superfamily: Aeropyrum pernix hypothetical protein APE0195

Query Match 76.5%; Score 13; DB 2; Length 121;  
Best Local Similarity 50.0%; Pred. No. 5.5e+02;  
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 2 EXXXLL 7

Db 34 ETASLL 39

## RESULT 12

S73426  
probable lipoprotein D09\_orf123 - Mycoplasma pneumoniae (strain ATCC 29342)  
C/Species: Mycoplasma pneumoniae  
A/Variety: ATCC 29342  
C/Date: 26-Feb-1997 #sequence\_revision 25-Apr-1997 #text\_change 23-Sep-2002  
R/Himmelreich, R.; Hilbert, H.; Plagens, H.; Pirkil, E.; Li, B.C.; Herrmann, R.  
Nucleic Acids Res. 24, 4420-4449, 1996  
A/Title: Complete sequence analysis of the genome of the bacterium Mycoplasma pneumoniae  
A/Reference number: S73327; MUID:97105885; PMID:8948633  
A/Accession: S73426  
A/Status: preliminary; nucleic acid sequence not shown; translation not shown  
A/Molecule type: DNA  
A/Residues: 1-123 <HIM>  
A/Cross-references: EMBL:AE000012; GB:U00089; NID:G1673755; PIDN:AAB95748.1; PID:G1673755  
Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1996  
A/Genetic code: SGC3  
C/Superfamily: hypothetical protein P02\_orf253

Query Match 76.5%; Score 13; DB 2; Length 123;  
Best Local Similarity 50.0%; Pred. No. 5.6e+02;  
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7  
| | |  
Db 68 ETASLL 73

## RESULT 13

B72669  
hypothetical protein APE0776 - Aeropyrum pernix (strain K1)  
C/Species: Aeropyrum pernix  
C/Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 20-Aug-1999  
A/Accession: B72669  
R/Kawarabayasi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takahawa, H.; Takamiya, M.; Maeda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K DNA Res. 6, 83-101, 1999  
A/Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyrum  
A/Reference number: A72450; MUID:99310339; PMID:10382966  
A/Accession: B72669  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-129 <KAW>  
A/Cross-references: DDBJ:AP000060; NID:G5104188; PIDN:BAA79754.1; PID:dl043540; PID:G5104188  
A/Experimental source: strain K1  
C/Genetics:  
A/Gene: APE0776

Query Match 76.5%; Score 13; DB 2; Length 129;  
Best Local Similarity 50.0%; Pred. No. 5.9e+02;  
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7  
| | |  
Db 78 ETASLL 83

## RESULT 14

S63331  
probable membrane protein YNR005c - yeast (Saccharomyces cerevisiae)  
N/Alternate names: hypothetical protein N2036  
C/Species: Saccharomyces cerevisiae  
C/Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 19-Apr-2002  
A/Accession: S63331  
R/Aert, R.; Verhaaselt, P.; Voet, M.; Volckaert, G.  
submitted to the Protein Sequence Database, April 1996  
A/Reference number: S62910

A/Accession: S63331  
A/Molecule type: DNA  
A/Residues: 1-134 <AER>  
A/Cross-references: EMBL:Z71620; NID:G1302476; PID:G1302477; GSPDB:GN00014; MIPS:YNR005c  
A/Experimental source: strain S288C  
C/Genetics:  
A/Gene: MIPS:YNR005c  
A/Cross-references: SGD:S0005288  
A/Map position: 14R  
C/Superfamily: Yeast probable membrane protein YNR005c  
C/Keywords: transmembrane protein  
F53-69/Domain: transmembrane #status predicted <TM1>  
F111-127/Domain: transmembrane #status predicted <TM2>

Query Match 76.5%; Score 13; DB 1; Length 134;  
Best Local Similarity 50.0%; Pred. No. 6.1e+02;  
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7  
| | |  
Db 7 ESSSLL 12

## RESULT 15

E36790  
hypothetical protein 40 - ictaluriid herpesvirus 1 (strain auburn 1)  
C/Species: ictaluriid herpesvirus 1  
A/Note: host ictaluriid punctatus (channel catfish)  
C/Date: 17-Aug-1992 #sequence\_revision 17-Aug-1992 #text\_change 08-Oct-1999  
C/Accession: E36790  
R/Davison, A.J.  
submitted to GenBank, January 1992  
A/Description: Channel catfish virus: a new type of herpesvirus.  
A/Reference number: A36804  
A/Accession: E36790  
A/Molecule type: DNA  
A/Residues: 1-136 <DAV>  
A/Cross-references: GB:M75136; NID:G331209; PIDN:AAA88143.1; PID:G331250  
R/Davison, A.J.  
Virology 186, 9-14, 1992  
A/Title: Channel catfish virus: a new type of herpesvirus.  
A/Reference number: A39447; MUID:92087490; PMID:1727613  
A/Contents: annotation  
A/Note: neither protein nor nucleic acid sequence is given  
C/Genetics:  
A/Gene: 40  
C/Superfamily: ictaluriid herpesvirus 1 hypothetical protein 40

Query Match 76.5%; Score 13; DB 2; Length 136;  
Best Local Similarity 50.0%; Pred. No. 6.2e+02;  
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7  
| | |  
Db 112 ESTSLL 117

## RESULT 16

T48326  
hypothetical protein F15A17.80 - Arabidopsis thaliana  
C/Species: Arabidopsis thaliana (mouse-ear cress)  
C/Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 20-Apr-2000  
C/Accession: T48326  
R/Bevan, M.; Terry, N.; Ardiles, W.; Buysaert, C.; Dasseville, R.; De Clerck, R.; ewes, H.W.; Rudd, S.; Lemcke, K.; Mayer, K.F.X.  
submitted to the Protein Sequence Database, April 2000  
A/Reference number: Z24491  
A/Accession: T48326  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-136 <BEV>  
A/Cross-references: EMBL:AL163002  
A/Experimental source: cultivar Columbia; BAC clone F15A17

## C:Genetics:

A:Map position: 5  
A:introns: 17/3, 67/3, 92/2  
A:Note: F15A17.80

Query Match 76.5%; Score 13; DB 2; Length 136;  
Best Local Similarity 50.0%; Pred. No. 6.2e+02;  
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7  
| |  
Db 117 ETSSL 122

## RESULT 17

S70820  
H+-transporting two-sector ATPase (EC 3.6.3.14) epsilon chain - red alga (Cyanidium caldwellii)  
C:Species: Chloroplast Cyanidium caldwellii  
C:Date: 19-Mar-1997 #sequence\_revision 25-Apr-1997 #text\_change 03-Jun-2002  
Accession: S70820; S36413  
Jetrzewski, M.; Zetsche, K.  
Plant Mol. Biol. 23, 67-76, 1993  
A:Title: Organization of plastid-encoded ATPase genes and flanking regions including homology  
A:Reference number: S39512; MUID:94033298; PMID:8219057  
A:Accession: S70820  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-138 <KOS>  
A:CROSS-references: EMBL:X66698; NID:G396522; PID:CAA47242.1; PID:G396526  
A:Note: the source is designated as Galdieria sulphuraria  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1992  
C:Genetics:

A:Gene: atpE  
A:Genome: Chloroplast  
C:Superfamily: H+-transporting ATP synthase epsilon chain  
C:Keywords: ATP biosynthesis; chloroplast; hydrolase; membrane-associated complex; thylakoid

Query Match 76.5%; Score 13; DB 2; Length 138;  
Best Local Similarity 50.0%; Pred. No. 6.3e+02;  
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7  
| |  
Db 98 ESSLL 103

## RESULT 18

S70899  
H+-tail protein [imported] - Escherichia coli (strain O157:H7, substrain RIMD 0509952)  
C:Species: Escherichia coli  
C:Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 03-Aug-2001  
Accession: H90899  
R:Hayashi, T.; Makino, K.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.  
Gasaawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.  
DNA Res. 8, 11-22, 2001  
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genomic  
A:Reference number: A99629; MUID:21156231; PMID:11258796  
A:Accession: H90899  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-140 <HAY>  
A:CROSS-references: GB:BA000007; PID:BA000007; PID:G13361634; GSPDB:GN00154  
A:Experimental source: strain O157:H7, substrain RIMD 0509952  
C:Genetics:

A:Gene: ECs2168  
C:Superfamily: phage lambda minor tail protein G

Query Match 76.5%; Score 13; DB 2; Length 140;  
Best Local Similarity 50.0%; Pred. No. 6.4e+02;  
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7  
| |

## Db

54 ETASLL 59

## RESULT 19

G90997  
Probable minor tail protein [imported] - Escherichia coli (strain O157:H7, substrain RU)  
C:Species: Escherichia coli  
C:Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 03-Aug-2001  
Accession: G90997  
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.  
Gasaawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.  
DNA Res. 8, 11-22, 2001  
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genomic  
A:Reference number: A99629; MUID:21156231; PMID:11258796  
A:Accession: G90997  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-140 <HAY>  
A:CROSS-references: GB:BA000007; PID:BA000007; PID:G13362420; GSPDB:GN00154  
A:Experimental source: strain O157:H7, substrain RIMD 0509952  
C:Genetics:

A:Gene: ECs2951  
C:Superfamily: phage lambda minor tail protein G

Query Match 76.5%; Score 13; DB 2; Length 140;  
Best Local Similarity 50.0%; Pred. No. 6.4e+02;  
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7  
| |  
Db 54 ETASLL 59

## RESULT 20

F85817  
Probable minor tail protein [imported] - Escherichia coli (strain O157:H7, substrain ED)  
C:Species: Escherichia coli  
C:Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 27-Nov-2001  
Accession: F85817  
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew,  
Miller, L.; Grobeck, B.J.; Davis, N.W.; Lim, A.; Diallanta, E.; Potamousis, K.; Apodaca,  
Nature 409, 529-533, 2001  
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.  
A:Reference number: A85480; MUID:21074935; PMID:11206551  
A:Accession: F85817  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-140 <STO>  
A:CROSS-references: GB:AB005174; NID:G12516103; PID:AA057002.1; GSPDB:GN00145; UWGP:Z3C  
A:Experimental source: strain O157:H7, substrain EDL933  
C:Genetics:

A:Gene: Z3086  
C:Superfamily: phage lambda minor tail protein G

Query Match 76.5%; Score 13; DB 2; Length 140;  
Best Local Similarity 50.0%; Pred. No. 6.4e+02;  
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7  
| |  
Db 54 ETASLL 59

## RESULT 21

H64361  
Hypothetical protein MJ0496 - Methanococcus jannaschii  
C:Species: Methanococcus jannaschii  
C:Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 21-Jul-2000  
Accession: H64361  
R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake,  
Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinscock, K.G.; Merrick, J.M.; Glodek, A.;  
raon, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.  
Science 273, 1058-1073, 1996

A;Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C.  
 A;Title: Complete genome sequence of the methanogenic archaeon, *Methanococcus jannaschii*  
 A;Reference number: A64300; MUID:196337999; PMID:868087  
 A;Accession: H64361  
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A;Molecule type: DNA  
 A;Residues: 1-141 <BUL>  
 A;Cross-references: GB:U67499; GB:L77117; NID:91591190; PIDN:AAB98489.1; PID:91591200; T  
 A;Genetics:  
 A;Map position: FOR438403-438828  
 A;Start codon: TTG

Query Match 76.5%; Score 13; DB 2; Length 141;  
 Best Local Similarity 50.0%; Pred. NO. 6.6e+02;  
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7  
 DB 42 ESAALL 47

LT 22  
 C70937  
 Hypothetical protein RV2185C - *Mycobacterium tuberculosis* (strain H37RV)  
 C;Species: *Mycobacterium tuberculosis*  
 C;Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 20-Jun-2000  
 C;Accession: C70937  
 R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.;  
 Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.;  
 Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.  
 Nature 393, 537-544, 1998

A;Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
 A;Title: Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome  
 A;Reference number: A70500; MUID:98295987; PMID:9634230  
 A;Accession: C70937  
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A;Molecule type: DNA  
 A;Residues: 1-144 <COL>  
 A;Cross-references: GB:AL021957; GB:AL123456; NID:93242293; PIDN:CAA17489.1; PID:9291110  
 A;Experimental source: strain H37RV  
 A;Genetics:  
 A;Gene: RV2185C

C;Superfamily: Streptomyces coelicolor hypothetical protein SC6G10.02c  
 Query Match 76.5%; Score 13; DB 2; Length 144;  
 Best Local Similarity 50.0%; Pred. NO. 6.6e+02;  
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7  
 DB 83 ESSSLL 88

RESULT 23  
 C87020  
 Conserved hypothetical protein ML0889 [imported] - *Mycobacterium leprae*  
 C;Species: *Mycobacterium leprae*  
 C;Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 10-May-2001  
 C;Accession: C87020  
 R;Cole, S.T.; Eigmeier, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.; Ho  
 sam, M.A.; Rutherford, K.M.  
 Nature 409, 1007-1011, 2001

A;Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.; S  
 A;Title: Massive gene decay in the leprosy bacillus.  
 A;Reference number: A86909; MUID:21128732; PMID:11234002  
 A;Accession: C87020  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-144 <STO>  
 A;Cross-references: GB:AL450380; NID:913092961; PIDN:CAC31270.1; GSPDB:GN00147  
 A;Genetics:  
 A;Gene: ML0889

C;Superfamily: Streptomyces coelicolor hypothetical protein SC6G10.02c

Query Match 76.5%; Score 13; DB 2; Length 144;  
 Best Local Similarity 50.0%; Pred. NO. 6.6e+02;  
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7  
 DB 83 ESSSLL 88

RESULT 24  
 A81824  
 Hypothetical protein NMA1955 [imported] - *Neisseria meningitidis* (strain Z249) serog  
 C;Species: *Neisseria meningitidis*  
 C;Date: 05-May-2000 #sequence\_revision 05-May-2000 #text\_change 02-Feb-2001  
 C;Accession: A81824  
 R;Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; R  
 Holroyd, S.; Jørgensen, K.; Leather, S.; Mouton, S.; Mungall, K.; Quail, M.A.; Rajandr  
 Nature 404, 502-506, 2000  
 A;Title: Complete DNA sequence of a serogroup A strain of *Neisseria meningitidis* Z249  
 A;Reference number: A81775; MUID:20222556; PMID:10761919

A;Accession: A81824  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-144 <PAR>  
 A;Cross-references: GB:AL162757; GB:AL457959; NID:97380371; PIDN:CAB85175.1; PID:973  
 A;Experimental source: serogroup A, strain Z2491  
 C;Genetics:  
 A;Gene: NMA1955

Query Match 76.5%; Score 13; DB 2; Length 144;  
 Best Local Similarity 50.0%; Pred. NO. 6.6e+02;  
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7  
 DB 24 EAASLL 29

RESULT 25  
 D70007  
 Hypothetical protein yueC - *Bacillus subtilis*  
 C;Species: *Bacillus subtilis*  
 C;Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 28-Jul-2000  
 C;Accession: D70007  
 R;Kunst, F.; Ogasawara, N.; Moser, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Be  
 C.; Bron, S.; Brouillet, S.; Brusa, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.;  
 A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabrat, C.; Ferrari,  
 Nature 390, 249-256, 1997

A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Ga  
 lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo,  
 Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardir  
 A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Maude, S.; Ma  
 Y.; M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portet  
 Rieser, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scan  
 A;Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; S  
 Akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiy  
 T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshid  
 A;Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Banchin, A.  
 A;Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtil*  
 A;Reference number: A69580; MUID:98044033; PMID:9384377

A;Accession: D70007  
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A;Molecule type: DNA  
 A;Residues: 1-151 <KUN>  
 A;Cross-references: GB:Z99120; GB:AL009126; NID:92635613; PIDN:CAB15173.1; PID:el184  
 A;Experimental source: strain 168  
 C;Genetics:  
 A;Gene: yueC

C;Superfamily: *Bacillus subtilis* hypothetical protein yueC  
 Query Match 76.5%; Score 13; DB 2; Length 151;

Best Local Similarity 50.0%; Pred. No. 6.9e+02;  
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7  
|  
|  
Db 115 ETSLL 120

RESULT 26  
E72487  
hypothetical protein APE2540 - Aeropyrum pernix (strain K1)  
C/Species: Aeropyrum pernix  
C/Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 20-Aug-1999  
C/Accession: E72487  
R;Kawarayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takah  
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K  
DNA Res. 6, 83-101, 1999  
A/Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr  
A/Reference number: A72450; MUID:99310339; PMID:10382966  
C/Accession: E72487  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-152 <KAW>  
A/Cross-references: DDBJ:AP000064; NID:95105945; PIDN:BAA81557.1; PID:d1045343; PID:g510  
A/Experimental source: strain K1  
C/Genetics:  
A/Gene: APE2540

Query Match 76.5%; Score 13; DB 2; Length 152;  
Best Local Similarity 50.0%; Pred. No. 6.9e+02;  
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7  
|  
|  
Db 30 EAAALL 35

RESULT 27  
C75438  
hypothetical protein - Deinococcus radiodurans (strain R1)  
C/Species: Deinococcus radiodurans  
C/Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 28-Jul-2000  
C/Accession: C75438  
R;White, O.P.; Eelsen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;  
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma  
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.  
Science 286, 1571-1577, 1999  
A/Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.  
A/Reference number: A75250; MUID:20036896; PMID:10567266  
C/Accession: C75438  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-161 <WHL>  
A/Cross-references: GB:AE001959; GB:AE000513; NID:96458820; PIDN:AAF10668.1; PID:g645882  
A/Experimental source: strain R1  
C/Genetics:  
A/Gene: DR1091  
A/Map position: 1  
C/Superfamily: Deinococcus radiodurans hypothetical protein DR1091

Query Match 76.5%; Score 13; DB 2; Length 161;  
Best Local Similarity 50.0%; Pred. No. 7.3e+02;  
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7  
|  
|  
Db 19 ETSLL 24

RESULT 28  
G70827  
hypothetical protein RV0459 - Mycobacterium tuberculosis (strain H37RV)  
C/Species: Mycobacterium tuberculosis

C/Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 20-Jun-2000  
C/Accession: G70827  
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.  
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.  
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.  
Nature 393, 537-544, 1998  
A/Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
A/Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
A/Reference number: A70500; MUID:98295987; PMID:9634230  
C/Accession: G70827  
A/Status: preliminary; nucleic acid sequence not shown; translation not shown  
A/Molecule type: DNA  
A/Residues: 1-163 <COL>  
A/Cross-references: GB:AL021933; GB:AL123456; NID:93261529; PIDN:CAA17414.1; PID:g29095;  
A/Experimental source: strain H37RV  
C/Genetics:  
A/Gene: RV0459  
C/Superfamily: Mycobacterium tuberculosis hypothetical protein RV0459

Query Match 76.5%; Score 13; DB 2; Length 163;  
Best Local Similarity 50.0%; Pred. No. 7.4e+02;  
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7  
|  
|  
Db 12 EAAALL 17

RESULT 29  
T36616  
hypothetical protein SCH35.43c - Streptomyces coelicolor  
C/Species: Streptomyces coelicolor  
C/Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 03-Dec-1999  
C/Accession: T36616  
R;Oliver, K.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.  
submitted to the EMBL Data Library, June 1999  
A/Reference number: 221610  
A/Accession: T36616  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-166 <OLI>  
A/Cross-references: EMBL:AL078610; PIDN:CAB44384.1; GSPDB:GN00070; SCOREDB:SCH35.43c  
A/Experimental source: strain A3(2)  
C/Genetics:  
A/Gene: SCOREDB:SCH35.43c

Query Match 76.5%; Score 13; DB 2; Length 166;  
Best Local Similarity 50.0%; Pred. No. 7.6e+02;  
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7  
|  
|  
Db 119 EAAALL 124

RESULT 30  
G86796  
Prophage p13 protein 01 [imported] - Lactococcus lactis subsp. lactis (strain IL1403)  
C/Species: Lactococcus lactis subsp. lactis  
C/Date: 23-Mar-2001 #sequence\_revision 23-Mar-2001 #text\_change 03-Aug-2001  
C/Accession: G86796  
R;Bolotin, A.; Winkler, P.; Mauger, S.; Jaillon, O.; Malarre, K.; Weissenbach, J.; Ehrli  
Genome Res. 11, 731-753, 2001  
A/Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis ss  
A/Reference number: A86625; MUID:21235186; PMID:11337471  
C/Accession: G86796  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-176 <STO>  
A/Cross-references: GB:AE005176; PID:gl2724360; PIDN:AAK05473.1; GSPDB:GN00146  
A/Experimental source: strain IL1403  
C/Genetics:  
A/Gene: p1301



R.; Davies, R.M.; Devlin, K.; Duthoy, S.; Feltwell, T.; Fraser, A.; Hamlin, N.; Holroyd, M.A.; Rutherford, K.M.  
 Nature 409, 1007-1011, 2001  
 A:Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.; Squares, R.  
 A:Title: Massive gene decay in the leprosy bacillus.  
 A:Reference number: A86909; MUID:21128732; PMID:11234002  
 A:Accession: B87138  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-181 <STO>  
 A:Cross-references: GB:AL450380; NID:g13093532; PIDN:CAC30786.1; GSPDB:GN00147  
 C:Genetics:  
 A:Gene: adk  
 C:Superfamily: adenylate kinase

Query Match 76.5%; Score 13; DB 2; Length 181;  
 Best Local Similarity 50.0%; Pred. No. 8.3e+02;  
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 QY 2 EXXXLL 7  
 DB 147 ETASLL 152

RESULT 36  
 S52661  
 ubiquitin-protein ligase (EC 6.3.2.19) - Arabidopsis thaliana  
 N:Alternate names: ubiquitin-conjugating enzyme E2; ubiquitin-conjugating protein UBC6  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 14-Jul-1995 #sequence\_revision 21-Jul-1995 #text\_change 03-Jun-2002  
 C:Accession: S52661; H84897  
 R:Watte, F.Z.; Butt, N.; Layfield, P.; Machuka, J.; Burke, J.F.; Moore, A.L.  
 Plant Mol. Biol. 26, 445-451, 1994  
 A:Title: Floral expression of a gene encoding an E2-relatedubiquitin-conjugating protein  
 A:Reference number: S52661; MUID:95036015; PMID:7948890  
 A:Accession: S52661  
 A:Molecule type: DNA  
 A:Residues: 1-183 <WAT>  
 R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.B.; Umayam, L.; Tallon, L.  
 Nature 402, 761-768, 1999  
 A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
 A:Reference number: A84420; MUID:20083487; PMID:10617197  
 A:Accession: H84897  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-183 <STO>  
 A:Cross-references: GB:AE002093; NID:g3702350; PIDN:AAC62907.1; GSPDB:GN00139  
 C:Genetics:  
 A:Gene: At2g46030  
 A:Map position: 2  
 C:Superfamily: yeast ubiquitin-protein ligase UBC1  
 C:Keywords: ligase

Query Match 76.5%; Score 13; DB 2; Length 183;  
 Best Local Similarity 50.0%; Pred. No. 8.3e+02;  
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 QY 2 EXXXLL 7  
 DB 123 EAASLL 128

RESULT 37  
 B72383  
 thymidine kinase (EC 2.7.1.21) [similarity] - Thermotoga maritima (strain MSB8)  
 C:Species: Thermotoga maritima  
 C:Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 31-Dec-2001  
 C:Accession: B72383  
 R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwin, M.L.; Dodson, R.J.; Haft, D.H.; Hickey, Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Cotton, M.S.; Phillips, C.A.; Richardson, D.; C.M.  
 Nature 399, 323-329, 1999  
 A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome sequence  
 A:Reference number: A72200; MUID:99287316; PMID:10360571  
 A:Accession: B72383  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-184 <ARN>  
 A:Cross-references: GB:AE001719; GB:AE000512; NID:g4980893; PIDN:AAD35486.1; PID:g4980893  
 A:Experimental source: strain MSB8  
 C:Genetics:  
 A:Gene: TM0401  
 C:Superfamily: thymidine kinase  
 C:Keywords: phosphotransferase

Query Match 76.5%; Score 13; DB 2; Length 184;  
 Best Local Similarity 50.0%; Pred. No. 8.4e+02;  
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 QY 2 EXXXLL 7  
 DB 123 EAAALL 128

RESULT 38  
 T50342  
 ubiquitin conjugating enzyme [imported] - fission yeast (Schizosaccharomyces pombe)

C:Species: Schizosaccharomyces pombe  
 C:Date: 09-Jun-2000 #sequence\_revision 09-Jun-2000 #text\_change 21-Jul-2000  
 C:Accession: T50342  
 R:Wood, V.; Rajandream, M.A.; Barrell, B.G.; Davis, P.; Churcher, C.M.  
 submitted to the EMBL Data Library, February 2000  
 A:Reference number: Z25063  
 A:Accession: T50342  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-184 <WOO>  
 A:Cross-references: EMBL:AL139314; PIDN:CAB75415.1; GSPDB:GN00067; SPDB:SPBC211.07C  
 A:Experimental source: strain 972h(-); cosmid c211  
 C:Genetics:  
 A:Gene: SPDB:SPBC211.07C  
 A:Map position: 2  
 C:Superfamily: yeast ubiquitin-protein ligase UBC1  
 C:Keywords: ligase

Query Match 76.5%; Score 13; DB 2; Length 184;  
 Best Local Similarity 50.0%; Pred. No. 8.4e+02;  
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 QY 2 EXXXLL 7  
 DB 123 EAAALL 128

RESULT 39  
 S43785  
 ubiquitin-protein ligase (EC 6.3.2.19) - Arabidopsis thaliana  
 N:Alternate names: ubiquitin-conjugating enzyme UBC6  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 27-Jan-1995 #sequence\_revision 27-Jan-1995 #text\_change 03-Jun-2002  
 C:Accession: S43785; S48897  
 R:Sullivan, M.L.; Carpenter, T.B.; Vierstra, R.D.  
 Plant Mol. Biol. 24, 651-661, 1994  
 A:Title: Homologues of wheat ubiquitin-conjugating enzymes - TaUBC1 and TaUBC4 are encoded  
 A:Reference number: S43781; MUID:94207190; PMID:8155884  
 A:Accession: S43785  
 A:Status: nucleic acid sequence not shown  
 A:Molecule type: DNA  
 A:Residues: 1-185 <SUL>  
 A:Cross-references: EMBL:LI9355  
 R:Sullivan, M.L.; Carpenter, T.B.; Vierstra, R.D.  
 submitted to the EMBL Data Library, December 1993  
 A:Description: Wheat ubiquitin conjugating enzymes Ta UBC1 and Ta UBC4 are encoded by sm  
 A:Reference number: S48896  
 A:Accession: S48897



A:Molecule type: DNA  
 A:Residues: 1-148, 151-185 <SUM>  
 A:Cross-references: EMBL:LI9355  
 C:Genetics:  
 A:Gene: UBC6  
 A:Introns: 15/2; 42/1; 80/2; 98/1; 141/1  
 C:Superfamily: yeast ubiquitin-protein ligase UBC1  
 C:Keywords: ligase

Query Match 76.5%; Score 13; DB 2; Length 185;  
 Best Local Similarity 50.0%; Pred. No. 8.4e+02;  
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7  
 | | |  
 DB 123 EAASLL 128

## RESULT 40

D97709  
 [Synthetic protein RC0076 [imported] - Rickettsia conorii (strain Malish 7)]  
 Species: Rickettsia conorii  
 C:Date: 30-Sep-2001 #sequence\_revision 30-Sep-2001 #text\_change 30-Sep-2001  
 C:Accession: D97709  
 R:Ogata, H.; Audic, S.; Renesto-Audiffren, P.; Fournier, P.E.; Barbe, V.; Samson, D.; Rickettsia conorii  
 Science 293, 2093-2098, 2001  
 A:Title: Mechanisms of Evolution in Rickettsia conorii and Rickettsia prowazekii.  
 A:Reference number: A97700; MUID:21442074; PMID:11557893  
 A:Accession: D97709  
 A:Status: Preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-186 <KUR>  
 A:Cross-references: GB:AE006914; PIDN:AAL02614.1; PID:G15619114; GSPDB:GN00173  
 C:Genetics:  
 A:Gene: RC0076

Query Match 76.5%; Score 13; DB 2; Length 186;  
 Best Local Similarity 50.0%; Pred. No. 8.5e+02;  
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7  
 | | |  
 DB 47 EASTLL 52

Search completed: November 5, 2003, 14:18:24  
 Job time : 22 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2003, 14:14:03 ; Search time 11 Seconds

(without alignments)  
29.926 Million cell updates/sec

Title: US-09-620-840C-2

Perfect score: 17  
Sequence: 1 XEXXXLL 7

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Matched: 127863 seqs, 47026705 residues

Number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	13	76.5	109	1	PRVA_RANES
2	13	76.5	110	1	PRVA_RANCA
3	13	76.5	115	1	Y025_BHP1
4	13	76.5	123	1	YAS4_MCPN
5	13	76.5	134	1	YN82_YEAST
6	13	76.5	136	1	VG40_HSV11
7	13	76.5	138	1	ATPE_GALSI
8	13	76.5	141	1	Y496_METJA
9	13	76.5	143	1	VAT1_MOUSE
10	13	76.5	158	1	NDQC_PSEFL
11	13	76.5	158	1	YV02_HUMAN
12	13	76.5	166	1	LAFL_VIBPA
13	13	76.5	167	1	LYCV_BPPS3
14	13	76.5	177	1	NR13_COTJA
15	13	76.5	181	1	KAD_MYCLE
16	13	76.5	183	1	UBC6_ARATH
17	13	76.5	184	1	KITH_THEMEA
18	13	76.5	193	1	NDQC_PSEAB
19	13	76.5	194	1	NDQC_PSEPU
20	13	76.5	195	1	FPD3_ARATH
21	13	76.5	198	1	FINW_SALTY
22	13	76.5	200	1	EPF2_CHICK
23	13	76.5	207	1	YL96_RALSO
24	13	76.5	208	1	VATE_CHLNP
25	13	76.5	211	1	YFDX_ECO57
26	13	76.5	211	1	YFDX_ECOLI
27	13	76.5	227	1	ATP6_ANAPL
28	13	76.5	227	1	ATP6_CHICK
29	13	76.5	227	1	ATP6_STRCA
30	13	76.5	234	1	NAGB_LISIN
31	13	76.5	234	1	NAGB_LISMO
32	13	76.5	238	1	SFSA_VIBPA
33	13	76.5	248	1	VMT1_INBAC

34	13	76.5	248	1	VMT1_INBAD	P13880 influenza b
35	13	76.5	248	1	VMT1_INBLE	P03489 influenza b
36	13	76.5	248	1	VMT1_INBSI	P06816 influenza b
37	13	76.5	252	1	UDP_ECOLI	P12758 escherichia
38	13	76.5	252	1	UDP_SALTY	O33808 salmonella
39	13	76.5	253	1	UDP_KLBAE	O08444 klebsiella
40	13	76.5	256	1	NAGE_MYCPU	Q98dj9 mycoplasma
41	13	76.5	261	1	OMPY_CHLMU	O9p223 chlamydia m
42	13	76.5	261	1	OMPY_CHLTP	O84376 chlamydia t
43	13	76.5	262	1	OMPY_CHLTP	Q926m5 chlamydia p
44	13	76.5	266	1	THID_ECOLI	P76422 escherichia
45	13	76.5	266	1	THID_SALTY	P55982 salmonella

## ALIGNMENTS

RESULT 1  
PRVA\_RANES STANDARD; PRT; 109 AA.  
ID AC P02627;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Parvalbumin alpha.  
OS Rana esculenta (edible frog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranioidea; Ranidae; Rana.  
OX NCBI\_TaxID=8401;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=82186701; PubMed=7042341;  
RA Jaurégui-Adell J., Pechere J.-F., Briand G., Richet C.,  
RA Demallie J.G.;  
RT "Amino-acid sequence of an alpha-parvalbumin, pI = 4.88, from frog  
skeletal muscle.";  
RL Eur. J. Biochem. 123:337-345(1982).  
RN [2]  
RP CALCIUM-BINDING DATA.  
RX MEDLINE=8610726; PubMed=3501318;  
RA Tanokura M., Yamada K.;  
RT "Heat capacity and entropy changes of the two major isotypes of  
bullfrog (Rana catesbeiana) parvalbumins induced by calcium  
binding.";  
RL Biochemistry 26:7668-7674(1987).  
CC -!- FUNCTION: IN MUSCLE, THE CALCIUM-BINDING PROTEIN PARVALBUMIN IS  
THOUGHT TO BE INVOLVED IN MUSCLE RELAXATION.  
CC -!- MISCELLANEOUS: THIS PARVALBUMIN HAS AN ISOELECTRIC POINT OF 4.88.  
CC -!- MISCELLANEOUS: THIS PROTEIN BINDS TWO CALCIUM IONS.  
CC -!- SIMILARITY: TO OTHER EF-HAND CALCIUM BINDING PROTEINS, BELONGS TO  
THE PARVALBUMINS SUBFAMILY.  
DR PIR; R03062; PVFGA.  
DR HSPF; P02625; LRTP.  
DR InterPro; IPR002048; EF-hand.  
DR Pfam; PF00036; ehand; 2.  
DR ProDom; PD000012; EF-hand; 1.  
DR SMART; SM00054; EFh; 2.  
DR PROSITE; PS00018; EF\_HAND; 2.  
KW Calcium-binding; Muscle protein; Repeat.  
FT CA BIND 51 62 EF-HAND 1 (BY SIMILARITY).  
FT CA BIND 90 101 EF-HAND 2 (BY SIMILARITY).  
SQ SEQUENCE 109 AA; 11797 MW; PD35BE18DDCT65C4 CRC64;

Query Match 76.5%; Score 13; DB 1; Length 109;  
Best Local Similarity 50.0%; Pred. No. 2.4e+02;  
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Oy 2 EXXXLL 7  
Db 81 ETALL 86

RESULT 2

```

PRVA_RANCA
ID PRVA_RANCA STANDARD; PRT; 110 AA.
AC P18087;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Parvalbumin alpha (P4.97)
OS Rana catesbeiana (Bull frog)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranioidea; Ranidae; Rana.
OX NCBI_TaxID=8400;
RN [1]
RP SEQUENCE.
RC TISSUE=Skeletal muscle;
RX MEDLINE=90346151; PubMed=2384163;
RA Sasaki T., Tanokura M., Asaoka K.;
RT "The complete amino acid sequence of bullfrog (Rana catesbeiana)
RL parvalbumin p14.97."
RL FEBS Lett. 268:249-251(1990).
CC -1- FUNCTION: IN MUSCLE, THE CALCIUM-BINDING PROTEIN PARVALBUMIN IS
CC -1- THOUGHT TO BE INVOLVED IN MUSCLE RELAXATION.
CC -1- MISCELLANEOUS: THIS PARVALBUMIN HAS AN ISOELECTRIC POINT OF 4.97.
CC -1- MISCELLANEOUS: THIS PROTEIN BINDS TWO CALCIUM IONS.
CC -1- SIMILARITY: TO OTHER EF-HAND CALCIUM BINDING PROTEINS, BELONGS TO
CC THE PARVALBUMINS SUBFAMILY.
DR PIR; S11054; S11054.
DR HSP; P02625; IRTF.
DR InterPro: IPR002048; EF-hand.
DR Pfam: PF00036; ehand; 2.
DR ProDom: PD000012; EF-hand; 1.
DR SMART: SM00054; EFh; 2.
DR PROSITE: PS00018; EF_HAND; 2.
KW Calcium-binding; Muscle protein; Repeat.
FT CA_BIND 52 63 EF-HAND 1 (BY SIMILARITY).
FT CA_BIND 91 102 EF-HAND 2 (BY SIMILARITY).
SQ SEQUENCE 110 AA; 11920 MW; D6C5710B65DFD1D4B CRC64;

Query Match 76.5%; Score 13; DB 1; Length 110;
Best Local Similarity 50.0%; Pred. No. 2.4e+02;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7
DB 82 ETALL 87

RESULT 3
Y025_BPHP1 STANDARD; PRT; 115 AA.
P51729;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical 13.1 kDa protein in LYS 3'region (ORF25).
OS Bacteriophage HPI.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae.
OX NCBI_TaxID=10690;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HP1C1;
RX MEDLINE=85128433; PubMed=6098523;
RA Benjamin R.C., Fitzmaurice W.P., Huang P.C., Scocca J.J.;
RT "Nucleotide sequence of cloned DNA segments of the Haemophilus
RL influenzae bacteriophage HPIc1."
RL Gene 31:173-185(1984).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=HP1C1;
RX MEDLINE=86279738; PubMed=8710508;
RA Exposito D., Fitzmaurice W.P., Benjamin R.C., Goodman S.D.,
RA Waldman A.S., Scocca J.J.;
RT "The complete nucleotide sequence of bacteriophage HPI DNA."
RL Nucleic Acids Res. 24:2360-2368(1996).

```

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
DR EMBL; U24159; AAB09212.1; -
DR PIR; S69533; S69533.
KW Hypothetical protein; Transmembrane.
FT TRANSMEM 10 28 POTENTIAL.
SQ SEQUENCE 115 AA; 13071 MW; 99B63FC525FD9CC2 CRC64;

Query Match 76.5%; Score 13; DB 1; Length 115;
Best Local Similarity 50.0%; Pred. No. 2.5e+02;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7
DB 35 ETALL 40

RESULT 4
YAS4_MYCPN STANDARD; PRT; 123 AA.
ID YAS4_MYCPN
AC P75060;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical lipoprotein MPN054 precursor (D09_orf123).
GN MPN054 OR MP100.
OS Mycoplasma pneumoniae.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2104;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 29342 / M129;
RX MEDLINE=97105885; PubMed=8948633;
RA Himmelsreich R., Hilbert H., Plagens H., Pirkel E., Li B.-C.,
RA Herrmann R.;
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
RL pneumoniae."
RL Nucleic Acids Res. 24:4420-4449(1996).
CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor
CC (Potential).
CC -1- SIMILARITY: BELONGS TO THE MG439 / MG440 FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
DR EMBL; AE000012; AAB95748.1; -
DR PIR; S73426; S73426.
DR InterPro: IPR001595; Lipoprotein_3.
DR Pfam: PF00938; Lipoprotein_3; 1.
DR ProDom: PD003276; Lipoprotein_3; 1.
DR PROSITE: PS00013; PROKAR_LIPOPROTEIN; 1.
KW Hypothetical protein; Lipoprotein; Membrane; Signal;
KW Complete proteome.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 123 HYPOTHETICAL LIPOPROTEIN MPN054.
FT LIPID 20 20 N-ACYL DIGLYCERIDE (POTENTIAL).
SQ SEQUENCE 123 AA; 14170 MW; 2780F8C46C5B1404 CRC64;

Query Match 76.5%; Score 13; DB 1; Length 123;
Best Local Similarity 50.0%; Pred. No. 2.7e+02;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

Qy 2 EXXXLL 7
Db 68 ETASLL 73

RESULT 5
YN82_YEAST STANDARD; PRT; 134 AA.
ID YN82_YEAST
AC P53717;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Hypothetical 14.7 kDa protein in RPC34-VPS27 intergenic region.
GN YNR05C OR N2036.
OS Saccharomycetes cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
CV NCBI_TaxID=4932;
} [1]
SEQUENCE FROM N.A.
RC STRAIN=S288C / FY1679;
RX MEDLINE=95208356; PubMed=7900425;
RA Verhasselt P., Aert R., Voet M., Volckaert G.;
RT "Twelve open reading frames revealed in the 23.6 kb segment flanking the centromere on the Saccharomyces cerevisiae chromosome XIV right arm.";
RT arm.";
RT RT
RL Yeast 10:1355-1361(1994).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL out-put of the European Bioinformatics Institute. There are no restrictions on use by non-profit institutions as long as its content is not modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.ebi.ac.uk/ or send an email to license@ebi.ac.uk).
CC -----
DR EMBL; X77195; NOT ANNOTATED_CDS.
DR EMBL; 271620; CAJ96261.1; -.
DR PIR; S63331; S63331.
DR SGD; S0005288; YNR005C.
DR KW Hypothetical protein; Transmembrane.
DR TRANSMEM 110 130 POTENTIAL.
DR SEQUENCE 134 AA; 14726 MW; CQED08B593982067 CRC64;

Query Match 76.5%; Score 13; DB 1; Length 134;
--est Local Similarity 50.0%; Pred. No. 2.9e+02;
atches 3; Conservative 0; Mismatches 3; Indels 0;

Qy 2 EXXXLL 7
Db 7 ESSSLL 12

RESULT 6
VG40_HSV11 STANDARD; PRT; 136 AA.
ID VG40_HSV11
AC Q00145;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 01-DEC-1992 (Rel. 24, Last annotation update)
DE Hypothetical gene 40 protein.
GN 40.
OS Ictalurid herpesvirus 1 (Channel catfish virus) (CCV).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Ictalurid Herpes-like viruses.
CV NCBI_TaxID=10401;
} [1]
SEQUENCE FROM N.A.
RC STRAIN=Auburn 1;
RX MEDLINE=92087490; PubMed=1727613;
RA Davison A.J.;
RT "Channel catfish virus: a new type of herpesvirus.";

```

```

Virology 186:9-14(1992).
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; M75136; AAA88143.1; --
PIR; E36790; E36790.
Hypothetical protein.
SEQUENCE 136 AA; 15171 MW; 1888DB4970AA899B CRC64;
-----
Query Match 76.5% Score 13; DB 1; Length 136;
Best Local Similarity 50.0%; Pred.No. 2.9e+02;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
-----
QY 2 EXXXLL 7
DB 112 ESTSL 117
-----
RESULT 7
ATPE GALSU STANDARD; PRT; 138 AA.
AC Q08808;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE ATP synthase epsilon chain (EC 3.6.3.14) (ATP synthase F1 sector
DE epsilon subunit).
GN ATPE.
OS Galdieria sulphuraria (Red alga).
OC Chloroplast.
OC Eukaryota; Rhodophyta; Bangiophyceae; Porphyridiales; Porphyridiaceae;
OC Galdieria.
OX NCBI_TaxID=130081;
RN [1]
SEQUENCE FROM N.A.
STRAIN=14-1-1 / Isolate 107.79/Goettingen;
MEDLINE=94033298; PubMed=8219057;
Kostrewa M., Zetsche K.;
"Organization of plastid-encoded ATPase genes and flanking regions
including homologues of infB and tsf in the thermophilic red alga
Galdieria sulphuraria.";
Plant Mol. Biol. 23:67-76(1993).
RL -1- FUNCTION: Produces ATP from ADP in the presence of a proton
gradient across the membrane.
CC -1- CATALYTIC ACTIVITY: ATP + H(2)O + H(+) (In) = ADP + phosphate +
H(+) (Out).
CC -1- SUBUNIT: F-type ATPases have 2 components, CF(1) - the catalytic
core - and CF(0) - the membrane proton channel. CF(1) has five
subunits: alpha(3), beta(3), gamma(1), delta(1), epsilon(1). CF(0)
has three main subunits: A, B and C.
CC -1- SUBCELLULAR LOCATION: Chloroplast thylakoid membrane.
CC -1- SIMILARITY: Belongs to the ATPase epsilon chain family.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; X66698; CAA47242.1; --
HAMAP; MF 00530; -- 1.
InterPro: IPR001469; ATPsyntc_DE.
Pfam; PF00401; ATP-syntc_DE_1.
Pfam; PF02823; ATP-syntc_DE_N; 1.
ProDom; PD000944; ATPsyntc_DE; 1.

```

```

DR TIGRFAMs: TIGR01216; ATP_synt_epsilon; 1.
KW Hydrolase; ATP synthetase; CF(1); Hydrogen ion transport; Thylakoid;
KW Membrane; Chloroplast.
SQ SEQUENCE 138 AA; 15057 MW; CBS4868BDCB8A62A CRC64;

Query Match 76.5%; Score 13; DB 1; Length 138;
Best Local Similarity 50.0%; Pred. No. 36+02;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7
DB 98 ESSLL 103

RESULT 8
ID Y496 METJA STANDARD; PRT; 141 AA.
AC Q57919;
DT 01-NOV-1997 (Rel. 35, Created)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
Hypothetical protein MJ0496.
GN MJ0496.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcaceae; Methanocaldococcus.
OX NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., FitzGerald L.M., Clayton R.A., Gocayne J.D.,
RA Kervatage A.R., Dougherty B.A., Tomb J.-P., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.P., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geochagen N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
RA Uterback T.R., Kelley J.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Cotton M.D., Roberts K.M., Smith H.O., Woese C.R., Venter J.C.;
RA Klenk H.-P., Frazer C.M., Smith H.O., Woese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
RL Science 273:1058-1073(1996).

-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).

-----
DR EMBL; U67499; AAB98489.1; ...
DR PIR; H64361; H64361.
DR TIGR; MJ0496;
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 141 AA; 15692 MW; DA284F11A8CEB443 CRC64;

Query Match 76.5%; Score 13; DB 1; Length 141;
Best Local Similarity 50.0%; Pred. No. 36+02;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7
DB 42 ESAALL 47

RESULT 9
ID V496 METJA STANDARD; PRT; 143 AA.
AC Q62465;
DT 01-NOV-1997 (Rel. 35, Created)
DT 16-OCT-2001 (Rel. 40, Last annotation update)

```

```

DE GN Synaptic vesicle membrane protein VAT-1 homolog (Fragment).
VAT1 OR VAT-1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98241025; PubMed=981869;
RA Hayess K., Kraft R., Sachsinger J., Janke J., Beckmann G., Rohde K.,
RA Jandrig B., Benndorf R.;
RA "Mammalian protein homologous to VAT-1 of Torpedo californica:
RT isolation from Ehrlich ascites tumor cells, biochemical
RT characterization, and organization of its gene.";
RL J. Cell. Biochem. 69:304-315(1998).
CC -1- SIMILARITY: BELONGS TO THE ZINC-CONTAINING ALCOHOL DEHYDROGENASE
CC FAMILY. QUINONE OXIDOREDUCTASE SUBFAMILY. STRONG, TO TORPEDO
CC CALIFORNICA VAT-1.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X95562; CAA64807.1; ALT_INIT.
CC HSSP; P28304; IQOR.
CC MGD; MGI:1349450; Vat1.
CC InterPro: IPR002085; Adh_zn_family.
CC InterPro: IPR002364; QOR_zeta_crystal.
CC Pfam: PF00107; ADH_zinc_N; 1.
CC PROSITE; PS01162; QOR_ZETA_CRYSTAL; 1.
KW Oxidoreductase; Zinc; Synapse; Membrane.
FT NON TER 143 143
SQ SEQUENCE 143 AA; 15203 MW; 6F2FES9F91D99CA CRC64;

Query Match 76.5%; Score 13; DB 1; Length 143;
Best Local Similarity 50.0%; Pred. No. 31e+02;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7
DB 72 EAALL 77

RESULT 10
ID NDOC PSEFL STANDARD; PRT; 158 AA.
AC Q07825;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Naphthalene 1,2-dioxygenase beta subunit (EC 1.14.12.12) (Naphthalene
DE 1,2-dioxygenase ISP beta) (Fragment).
GN NDOC OR NDOC3.
OS Pseudomonas fluorescens.
OG Plasmid.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=294;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 17483;
RA Hamann C.;
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: COMPONENT OF NAPHTHALENE DIOXYGENASE (NDO)
CC MULTICOMPONENT ENZYME SYSTEM WHICH CATALYZES THE INCORPORATION OF
CC BOTH ATOMS OF MOLECULAR OXYGEN INTO NAPHTHALENE TO FORM CIS-
CC NAPHTHALENE DIHYDRODIOL. THE BETA SUBUNIT MAY BE RESPONSIBLE FOR
CC THE SUBSTRATE SPECIFICITY OF THE ENZYME.
CC -1- CATALYTIC ACTIVITY: Naphthalene + NADH + O(2) -> (1R,2S)-1,2-

```

CC dilynonaphthalene-1,2-diol + NAD(+) .

CC -1- COPACTOR: PROBABLY BINDS A 2FE-2S GROUP AND AN IRON ATOM.

CC -1- PATHWAY: UPPER NAPHTHALENE CATABOLIC PATHWAY WHICH INVOLVES

CC CONVERSION OF NAPHTHALENE TO SALICYLATE, AND CATABOLISM OF

CC DIBENZOXYTHIOPHENE (DBT) AND PHENANTHRENE. CONVERTS PHENANTHRENE TO

CC 1-HYDROXY-2-NAPHTHOIC ACID AND THE METABOLISM OF DBT IS LIMITED TO

CC OXIDATION OF THE AROMATIC RING.

CC -1- SYSTEM: NAPHTHALENE DIOXYGENASE (NDO) MULTICOMPONENT ENZYME

CC SUBUNIT IS COMPOSED OF AN ELECTRON TRANSFER COMPONENT AND AN IRON

CC SULFUR PROTEIN (ISP). THE ELECTRON TRANSFER COMPONENT IS COMPOSED

CC OF FERREDOXIN REDUCTASE (NDOR) AND FERREDOXIN (NDOA), AND ISP IS

CC COMPOSED OF A LARGE ALPHA SUBUNIT (NDOB) AND A SMALL BETA SUBUNIT

CC (NDOC).

CC -1- SIMILARITY: BELONGS TO THE BACTERIAL RING-HYDROXYLATING

CC DIOXYGENASE BETA SUBUNIT FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>)

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).

CC -----

CC EMBL: AF004283; AAB61371.1; -

CC HSPSP; P23095; NDO.

CC InterPro: IPR000391; Ring hydroxyl B.

CC Pfam: PF00866; Ring hydroxyl B; 1.

CC ProDom: PD002199; Ring hydroxyl B; 1.

CC KW Aromatic hydrocarbons Catabolism, NAD; Oxidoreductase; Dioxxygenase;

CC FT NON TER.

CC SQ SEQUENCE 158 AA; 18597 MW; 9C80BB72C22E7BF9 CRC64;

Query Match 76.5%; Score 13; DB 1; Length 158;

Best Local Similarity 50.0%; Pred. No. 3.4e+02;

Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7

DB 33 EATLL 38

RESULT 11

ID YV02\_HUMAN STANDARD; PRT; 158 AA.

PA Q99765;

30-MAY-2000 (Rel. 39, Created)

30-MAY-2000 (Rel. 39, Last sequence update)

28-FEB-2003 (Rel. 41, Last annotation update)

DE Hypothetical protein DJ1104E15.3.

OS Homo sapiens (human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Brain;

RX MEDLINE=97264341; PubMed=9110174;

RA Yu W., Anderson B., Worley K.C., Muzny D.M., Ding Y., Liu W.,

RA Ricafrente J.Y., Wentland M.A., Lennan G., Gibbs R.A.;

RT "Large-scale concatenation cDNA sequencing.";

RL Genome Res. 7:353-358(1997).

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE=20057165; PubMed=10591208;

RA Dunham I., Hunt A.R., Collins J.E., Bruskiewicz R., Beare D.M.,

RA Clamp M., Smink L.J., Ainscough R., Almeida J.P., Babbage A.K.,

RA Baguley C., Bailey J., Barlow K.F., Bates K.N., Beasley O.P.,

RA Bird C.P., Blakey S.E., Bridgeman A.M., Buck D., Burgess J.,

RA Burrill W.D., Burton J., Carder C., Carter N.P., Chen Y., Clark G.,

RA Clegg S.M., Cobley V.E., Cole C.G., Collier R.E., Connor R.,

RA Conroy D., Corby N.R., Coville G.J., Cox A.V., Davis J., Dawson E.,

RA Dhami P.D., Dockree C., Dodsworth S.J., Durbin R.M., Ellington A.G.,

RA Evans K.L., Fey J.M., Fleming K., French L., Garner A.A.,

RA Gilbert J.G.R., Goward M.E., Grafham D.V., Griffiths M.N.D., Hall C.,

RA Hall R.E., Hall-Tamlyn G., Heathcote R.W., Ho S., Holmes S.,

RA Hunt S.E., Jones M.C., Kerehaw J., Kimberley A.M., King A.,

RA Laird G.K., Langford C.P., Levarsha M.A., Lloyd C., Lloyd D.M.,

RA Martyn I.D., Mashregi-Mohammadi M., Matthews L.H., Mccann O.T.,

RA Mcclay J., McLaren S., McMurray A.A., Milne S.A., Mortimore B.J.,

RA Odell C.N., Pavitt R., Pearce A.V., Pearson D., Phillimore B.J.C.T.,

RA Phillips S.H., Plumb R.W., Ramsay H., Ramsey Y., Rogers L., Ross M.T.,

RA Scott C.E., Sehra H.K., Skuce C.D., Smalley S., Smith M.L.,

RA Soderlund C., Spragon L., Steward C.A., Sulston J.E., Swann R.M.,

RA Vaudin M., Wall M., Wallis J.M., Whiteley M.N., Willey D.B.,

RA Williams L., Williams S.A., Williamson H., Wilmer T.E., Wilming L.,

RA Wright C.L., Hubbard T., Bentley D.R., Beck S., Rogers J., Shimizu N.,

RA Minoshima S., Kawasaki K., Sasaki T., Asakawa S., Kudoh J.,

RA Shintani A., Shibuya K., Yoshizaki Y., Aoki N., Mitsuyama S.,

RA Roe B.A., Chen F., Chu L., Crabtree J., Deschamps S., Do A., Do T.,

RA Dorman A., Fang E., Fu Y., Hu P., Hua A., Kenton S., Lai H., Lao H.I.,

RA Lewis J., Lewis S., Lin S.-P., Loh P., Mala J.E., Nguyen T., Pan H.,

RA Phan S., Qi S., Qian Y., Ray L., Ren Q., Shaull S., Sloan D., Song L.,

RA Wang Q., Wang Y., Wang Z., White J., Willingham D., Wu H., Yao Z.,

RA Zhan M., Zhang G., Chissee S., Murray J., Miller N., Minx P.,

RA Fulton R., Johnson D., Bemis G., Bentley D., Bradshaw H., Bourne S.,

RA Cordes M., Du Z., Fulton L., Goela D., Graves T., Hawkins J.,

RA Hinds K., Kemp K., Matreille P., Layman D., Ozersky P., Rohlfing T.,

RA Scheet P., Walker C., Wansley A., Wohldmann P., Pepin K., Nelson J.,

RA Korf I., Bedell J.A., Hillier L., Mardis E., Waterston R., Wilson R.,

RA Emanuel B.S., Shaikh T., Kurahashi H., Saitta S., Budarf M.L.,

RA McDermid H.E., Johnson A., Wong A.C.C., Morrow B.E., Edelmann L.,

RA Kim U.J., Shizuya H., Simon M.I., Dumanaki J.P., Feyrard M., Kedra D.,

RA Seroussi E., Fransson I., Tapia I., Bruder C.E., O'Brien K.P.,

RA Wilkinson P., Bodenteich A., Hartman K., Hu X., Khan A.S., Lane L.,

RA Tilahun Y., Wright H.;

RT "The DNA sequence of human chromosome 22.";

RL Nature 402:489-495(1999).

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>)

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).

CC -----

CC EMBL: U79252; AAB50201.1; -

CC EMBL: AL022312; CAB45283.1; -

CC KW Hypothetical protein.

CC SQ SEQUENCE 158 AA; 17717 MW; A9DECCF7238B418 CRC64;

Query Match 76.5%; Score 13; DB 1; Length 158;

Best Local Similarity 50.0%; Pred. No. 3.4e+02;

Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7

DB 74 EAAALL 79

RESULT 12

LAFL\_VIBPA STANDARD; PRT; 166 AA.

ID LAFL\_VIBPA

AC Q03476; P74953;

DT 01-OCT-1993 (Rel. 27, Created)

DT 15-SEP-2003 (Rel. 42, Last sequence update)

DE Flagellar protein lafl.

OS LAFL OR LAFF OR VPA1554.

GN Vibrio parahaemolyticus.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;

OC Vibrionaceae; Vibrrio.

OX NCBI\_TaxID=570;

RN [1]

RT SEQUENCE FROM N.A.  
RC STRAIN=BB22;  
RA McCarter L.L., Noack D.;  
RT "The lateral flagellar hierarchy of gene expression.";  
RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=RMD 2210633 / Serotype O3:K6;  
RX MEDLINE=22508454; PubMed=12620739;  
RA Makino K., Oshima K., Kurokawa K., Yokoyama K., Uda T., Tagomori K.,  
RA Iijima Y., Najima M., Nakano M., Yamashita A., Kubota Y., Kimura S.,  
RA Yasunaga T., Honda T., Shinagawa H., Hattori M., Iida T.;  
RT "Genome sequence of *Vibrio parahaemolyticus*: a pathogenic mechanism  
RT distinct from that of *V. cholerae*.";  
RL Lancet 361:743-749 (2003).  
RN [3]  
RP SEQUENCE OF 93-166 FROM N.A.  
RC STRAIN=BB22;  
RX MEDLINE=93273702; PubMed=8501040;  
RA McCarter L.L., Wright M.E.;  
RT "Identification of genes encoding components of the swarmer cell  
RT flagellar motor and propeller and a sigma factor controlling  
RT differentiation of *Vibrio parahaemolyticus*.";  
RL J. Bacteriol. 175:3361-3371 (1993).  
CC -|- FUNCTION: CONTROLS THE ROTATIONAL DIRECTION OF FLAGELLA DURING  
CC CHEMOTAXIS.  
CC -|- INDUCTION: Under conditions in which the polar flagellum is not  
CC functional.  
CC -|- SIMILARITY: BELONGS TO THE FLIL FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; U52957; AAB07355.1; -  
DR EMBL; AF005089; BAC62897.1; -  
DR EMBL; U20541; AAA62352.1; -  
DR PIR; G40590; G40590.  
DR InterPro; IPR005503; FlilL.  
DR Pfam; PF03748; FlilL; 1  
KW Chemotaxis; Flagell; 1  
FT CONFLICT 93 94 TR -> AG (IN REF. 3).  
SQ SEQUENCE 166 AA; 18665 MW; B892CC8B7E866CD9 CRC64;  
Query Match 76.5%; Score 13; DB 1; Length 166;  
Best Local Similarity 50.0%; Pred. No. 3.5e+02;  
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 2 EXXXLL 7  
DB 41 ETSSL 46  
RESULT 13  
LYCV\_BPFS3  
ID LYCV\_BPFS3 STANDARD; PRT; 167 AA.  
AC O80288;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Lysozyme (EC 3.2.1.17) (Lysis protein) (Muramidase) (Endolysin)  
DE (Protein gp19).  
GN 19.  
OS Bacteriophage PS34.  
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Podoviridae.  
ON NCBI\_TaxID=83127;  
RP SEQUENCE FROM N.A.  
RA Zimmer A., Schmieger H.;

"Lysis gene modules in the phage P22 gene pool.";  
Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.  
-|- FUNCTION: ESSENTIAL FOR LYSIS OF BACTERIAL CELL WALL, BY SHOWING  
CELL WALL HYDROLYZING ACTIVITY (BY SIMILARITY).  
-|- CATALYTIC ACTIVITY: Hydrolysis of the 1,4-beta-linkages between N-  
acetyl-D-glucosamine and N-acetylmuramic acid in peptidoglycan  
heteropolymers of the prokaryotes cell walls.  
-|- SIMILARITY: BELONGS TO FAMILY 24 OF GLYCOSYL HYDROLASES.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; AJ011580; CAA09706.1; -  
DR InterPro; IPR002196; Glyco\_hydro\_24.  
DR Pfam; PF00959; Phage\_lysozyme; 1.  
KW Hydrolase; Glycosidase; Bacteriolytic enzyme.  
FT ACT\_SITE 37 37 PROTON DONOR (BY SIMILARITY).  
FT ACT\_SITE 46 46 NUCLEOPHILE (BY SIMILARITY).  
SQ SEQUENCE 167 AA; 18187 MW; B9DC214DA0382343 CRC64;  
Query Match 76.5%; Score 13; DB 1; Length 167;  
Best Local Similarity 50.0%; Pred. No. 3.6e+02;  
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 2 EXXXLL 7  
DB 117 ETSSL 122  
RESULT 14  
NR13\_COTJA  
ID NR13\_COTJA STANDARD; PRT; 177 AA.  
AC Q90343;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Apoptosis regulator NR-13.  
GN NR-13.  
OS Coturnix coturnix japonica (Japanese quail).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
OC Coturnix.  
ON NCBI\_TaxID=93934;  
RP SEQUENCE FROM N.A.  
RX TISSUE=Fibroblast, and Neuroretina;  
RX MEDLINE=95246730; PubMed=7729415;  
RA Gillet G., Guerin M., Trembleau A., Brun G.;  
RT "A Bcl-2-related gene is activated in avian cells transformed by the  
RT Rous sarcoma virus.";  
RL EMBO J. 14:1372-1381 (1995).  
CC -|- FUNCTION: ROUS SARCOMA VIRUS-ACTIVATED PROTEIN WITH ANTI-APOPTOTIC  
CC PROPERTIES.  
CC -|- SUBCELLULAR LOCATION: Plasma membrane.  
CC -|- TISSUE SPECIFICITY: MAINLY EXPRESSED IN NEURAL AND MUSCULAR  
CC TISSUES.  
CC -|- DEVELOPMENTAL STAGE: EXPRESSION IS DRAMATICALLY DOWN-REGULATED  
CC AFTER EMBRYONIC DAY 7 IN THE OPTIC TECTUM, AND CORRELATES WITH THE  
CC ONSET OF APOPTOSIS IN THIS AREA.  
CC -|- SIMILARITY: Contains 1 Bcl-2 homology 1 (BH1) domain.  
CC -|- SIMILARITY: Contains 1 Bcl-2 homology 2 (BH2) domain.  
CC -|- SIMILARITY: BELONGS TO THE BCL-2 FAMILY.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; X84418; CRA59136.1; -  
 CC HSSP; Q07817; INAZ.  
 CC InterPro; IPR000712; BCL2 BH.  
 CC InterPro; IPR002475; BCL2\_family.  
 CC Pfam; PF00452; BCL-2; 1.  
 CC SMART; SM00337; BCL; 1.  
 CC PROSITE; PS01080; BH1; 1.  
 CC PROSITE; PS01258; BH2; 1.  
 CC PROSITE; PS00662; BCL2\_FAMILY; 1.  
 CC Apoptosis; Transmembrane.  
 CC DOMAIN 75 94 BH1.  
 CC DOMAIN 126 141 BH2.  
 CC TRANSMEM 86 106 POTENTIAL.  
 CC TRANSMEM 156 177 POTENTIAL.  
 CC SEQUENCE 177 AA; 18826 MW; -AD5CE79D3353CC1P CRC64;

Query Match 76.5%; Score 13; DB 1; Length 177;  
 Best Local Similarity 50.0%; Pred. No. 3.8e+02;  
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7  
 DB 63 EAAALL 68

## RESULT 15

KAD MYCLE  
 ID KAD MYCLE STANDARD; PRT; 181 AA.  
 AC Q33007;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Adenylate kinase (EC 2.7.4.3) (ATP-AMP transphosphorylase).  
 GN ADK OR ML1832 OR MLCB2492.28.  
 OS Mycobacterium leprae.  
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
 OC NCBI\_TaxID=1769;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=TN;  
 RX MEDLINE=21128732; PubMed=11234002;  
 PA Cole S.T., Eigmeier K., Parkhill J., James K.D., Thomson N.R.,  
 Wheeler P.R., Honore N., Garnier T., Churcher C., Harris D.,  
 Mungall K., Basham D., Brown D., Chillingworth T., Connor R.,  
 Davies R.M., Devlin K., Duthoy S., Feltwell T., Fraser A., Hamlin N.,  
 Holtroyd S., Hornsby T., Jagels K., Lacroix C., Maclean J., Moule S.,  
 Murphy L., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,  
 Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,  
 Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,  
 RA Bartell B.G.;  
 RT "Massive gene decay in the leprosy bacillus";  
 RL Nature 409:1007-1011(2001).  
 CC -!- FUNCTION: THIS SMALL UBIQUITOUS ENZYME IS ESSENTIAL FOR  
 CC MAINTENANCE AND CELL GROWTH.  
 CC -!- CATALYTIC ACTIVITY: ATP + AMP = ADP + ADP.  
 CC -!- SUBUNIT: Monomer (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -!- SIMILARITY: Belongs to the adenylate kinase family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; 298756; CAB11460.1; -  
 CC EMBL; AL583923; CAC30786.1; -

DR PIR; B87138; B87138.  
 DR PIR; T45390; T45390.  
 DR HSSP; P27142; IZIN.  
 DR Leproma; ML1832; -  
 DR HAMAP; MF\_00235; -; 1.  
 DR InterPro; IPR006259; Adenyl\_kin\_sub.  
 DR InterPro; IPR000850; Adenylate\_Kin.  
 DR Pfam; PF00406; ADK; 1.  
 DR PRINTS; PR00094; ADENYLKINASE.  
 DR ProDom; PD000657; Adenylate\_Kin; 1.  
 DR TIGRFAMs; TIGR01351; adk; 1.  
 DR PROSITE; PS00113; ADENYLATE\_KINASE; 1.  
 DR TRANSFERASE; Kinase; ATP-binding; Complete proteome.  
 KW NP BIND 149 149 A -> R (IN REF. 1; CAB11460).  
 FT NP BIND 149 149 ATP (BY SIMILARITY)  
 FT CONFLICT 149 149 MW; 3023ASEFC359AG14 CRC64;  
 SQ SEQUENCE 181 AA; 20111 MW; 76.5%; Score 13; DB 1; Length 181;

Query Match 76.5%; Score 13; DB 1; Length 181;  
 Best Local Similarity 50.0%; Pred. No. 3.9e+02;  
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7  
 DB 147 ETASLL 152

## RESULT 16

UBC6\_ARATH  
 ID UBC6\_ARATH STANDARD; PRT; 183 AA.  
 AC P42750; O49739; O82365;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Ubiquitin-conjugating enzyme E2-21 kDa 3 (EC 6.3.2.19) (Ubiquitin-  
 DE protein ligase 6) (Ubiquitin carrier protein 6).  
 GN UBC6 OR AT2G46030 OR T3F17.32.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eumids II; Brassicales; Brassicaceae; Arabidopsids.  
 OC NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia; TISSUE=Green leaf;  
 RX MEDLINE=94207190; PubMed=8155884;  
 RA Sullivan M.L.; Carpenter T.B.; Vierstra R.D.;  
 RT "Homologues of wheat ubiquitin-conjugating enzymes -- TaUBC1 and  
 RT TaUBC4 are encoded by small multigene families in Arabidopsis  
 RT thaliana";  
 RL Plant Mol. Biol. 24:651-661(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=95036015; PubMed=7948890;  
 RA Watts P.Z.; Butt N.; Layfield P.; Machuka J.; Burke J.P.; Moore A.L.;  
 RT "Floral expression of a gene encoding an E2-related ubiquitin-  
 RT conjugating protein from Arabidopsis thaliana";  
 RL Plant Mol. Biol. 26:445-451(1994).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia;  
 RX MEDLINE=20083487; PubMed=10617197;  
 RA Lin X.; Kaul S.; Rounsley S.D.; Shea T.P.; Benito M.-I.; Town C.D.,  
 RA Fujii C.Y.; Mason T.M.; Bowman C.L.; Barnstead M.E.; Feldblum T.V.,  
 RA Buell C.R.; Ketchum K.A.; Lee J.J.; Rensing C.M.; Koo H.L.,  
 RA Moffat K.S.; Cronin L.A.; Shen M.; Van Aken S.; Umayam L.,  
 RA Tallon L.J.; Gill J.B.; Adams M.D.; Carrera A.J.; Creasy T.H.,  
 RA Goodman H.M.; Somerville C.R.; Copenhaver G.P.; Preuss D.,  
 RA Nierman W.C.; White O.; Eisen J.A.; Salzberg S.L.; Fraser C.M.,  
 RA Venter J.C.;  
 RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis  
 RT thaliana";  
 RL Nature 402:761-768(1999).  
 CC -!- FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO



OTHER PROTEINS.

-1- CATALYTIC ACTIVITY: ATP + ubiquitin + protein lysine = AMP + diphosphate + protein N-ubiquityllysine.

-1- PATHWAY: Ubiquitin conjugation, second step.

-1- TISSUE SPECIFICITY: EXPRESSED IN ALL TISSUES EXAMINED.

-1- MISCELLANEOUS: A cysteine residue is required for ubiquitin-thiolester formation.

-1- SIMILARITY: Belongs to the ubiquitin-conjugating enzyme family.

-1- CAUTION: Ref.2 sequence differs from that shown due to erroneous gene model prediction.

-----

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announcement/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

EMBL; L19355; AAA32901.1; -  
EMBL; X71381; CAA50503.1; ALT SEQ.  
EMBL; AC005397; AAC62907.1; -  
PIR; S43785; S43785.  
PIR; S52661; S52661.  
KSP; F15731; IQCQ.  
InterPro; IPR000608; UBQ conjugat.  
Pfam; PF00179; UQ con; 1.  
ProDom; PD000461; UBQ conjugat; 1.  
SMART; SM00212; UBQ; 1.  
PROSITE; PS00183; UBIQUITIN\_CONJUGAT\_1; 1.  
PROSITE; PS00127; UBIQUITIN\_CONJUGAT\_2; 1.  
KW Ub1 conjugation pathway; Ligase; Multigene family.  
FT BINDING '85 85 UBIQUITIN (BY SIMILARITY).  
CONFLICT 2 2 A -> S (IN REF. 1).  
SEQUENCE 183 AA; 20785 MW; D033BA90955DEDF2 CRC64;

Query Match 76.5%; Score 13; DB 1; Length 183;  
Best Local Similarity 50.0%; Pred. No. 3.9e+02;  
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7  
DB 123 EASLL 128

RESULT 17  
KITH THEMA STANDARD; PRT; 184 AA.  
IN QWY2;  
30-MAY-2000 (Rel. 39, Created)  
30-MAY-2000 (Rel. 39, Last sequence update)  
28-FEB-2003 (Rel. 41, Last annotation update)  
Thymidine kinase (EC 2.7.1.21).  
TKD OR TM0401.  
OS Thermotoga maritima.  
OC Bacteria; Thermotogae; Thermotogales; Thermotogaceae; Thermotoga.  
OX NCBI\_TaxId=2336;  
RN [1]  
SEQUENCE FROM N.A.  
STRAIN=MSB8 / DSM 3109;  
MEDLINE=99287316; PubMed=10360571;  
RA Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A., McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M., Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D., Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O., Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;  
RT "Evidence for lateral gene transfer between Archaea and Bacteria from genome sequence of Thermotoga maritima."  
RL Nature 393:323-329 (1999).  
CC -1- CATALYTIC ACTIVITY: ATP + thymidine = ADP + thymidine 5'-phosphate.  
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).

-1- SIMILARITY: BELONGS TO THE THYMIDINE KINASE FAMILY.

-----

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announcement/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

EMBL; AE001719; AAD35486.1; -  
PIR; B72383; B72383.  
TIGR; TM0401; -  
HAWAP; MF 00124; -; 1.  
InterPro; IPR001267; TK\_cell.  
Pfam; PF00265; TK; 1.  
PROSITE; PS00603; TK\_CELLULAR\_TYPE; 1.  
KW Transferase; Kinase; DNA synthesis; ATP-binding; Complete proteome.  
FT NP\_BIND 10 17 ATP (BY SIMILARITY).  
SEQUENCE 184 AA; 20654 MW; 906B2C1BE3A7EDDA CRC64;

Query Match 76.5%; Score 13; DB 1; Length 184;  
Best Local Similarity 50.0%; Pred. No. 3.9e+02;  
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7  
DB 121 ETALL 126

RESULT 18  
NDOC\_PSEAE STANDARD; PRT; 193 AA.  
ID NDOC\_PSEAE  
AC 051435;  
DT 15-DEC-1998 (Rel. 37, Created)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Naphthalene 1,2-dioxygenase beta subunit (EC 1.14.12.12) (Naphthalene 1,2-dioxygenase ISP beta).  
GN NDOC OR PAHA4.  
OS Pseudomonas aeruginosa.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
OC Pseudomonadaceae; Pseudomonas.  
OX NCBI\_TaxId=287;  
RN [1]  
SEQUENCE FROM N.A.  
STRAIN=PAK1;  
RA Takizawa N., Iida T., Yamauchi K., Satoh S., Wang Y., Fukuda M., Kiyohara H.;  
RT "The molecular analysis of an NAH7-type gene cluster, pah, located on the chromosome of Pseudomonas aeruginosa PAK1,"  
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: COMPONENT OF NAPHTHALENE DIOXYGENASE (NDO).  
CC MULTICOMPONENT ENZYME SYSTEM WHICH CATALYZES THE INCORPORATION OF BOTH ATOMS OF MOLECULAR OXYGEN INTO NAPHTHALENE TO FORM CIS-NAPHTHALENE DIHYDRODIOL. THE BETA SUBUNIT MAY BE RESPONSIBLE FOR THE SUBSTRATE SPECIFICITY OF THE ENZYME.  
CC -1- CATALYTIC ACTIVITY: Naphthalene + NADH + O(2) = (1R,2S)-1,2-dihydronaphthalene-1,2-diol + NAD(+).  
CC -1- COFACTOR: PROBABLY BINDS A 2FE-2S GROUP AND AN IRON ATOM.  
CC -1- PATHWAY: UPPER NAPHTHALENE CATABOLIC PATHWAY WHICH INVOLVES CONVERSION OF NAPHTHALENE TO SALICYLATE, AND CATABOLISM OF DIBENZOXYTHIOPHENE (DBT) AND PHENANTHRENE. CONVERTS PHENANTHRENE TO 1-HYDROXY-2-NAPHTHOIC ACID AND PHENANTHRENE. THE METABOLISM OF DBT IS LIMITED TO OXIDATION OF THE AROMATIC RING.  
CC -1- SUBUNIT: NAPHTHALENE DIOXYGENASE (NDO) MULTICOMPONENT ENZYME SYSTEM IS COMPOSED OF AN ELECTRON TRANSFER COMPONENT AND AN IRON SULFUR PROTEIN (ISP). THE ELECTRON TRANSFER COMPONENT IS COMPOSED OF FERREDOXIN REDUCTASE (NDOR) AND FERREDOXIN (NDOA), AND ISP IS COMPOSED OF A LARGE ALPHA SUBUNIT (NDOB) AND A SMALL BETA SUBUNIT (NDOC).  
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL RING-HYDROXYLATING DIOXYGENASE BETA SUBUNIT FAMILY.

-----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----  
 ENBL; D84146; BAA12241.1; -  
 HSPB; P23095; INDO.  
 InterPro: IPR00391; Ring\_hydroxyl\_B.  
 Pfam: PF00866; Ring\_hydroxyl\_B; 1.  
 ProDom: PD002199; Ring\_hydroxyl\_B; 1.  
 KW Aromatic hydrocarbons catabolism; Oxidoreductase; NAD; Dioxigenase.  
 SQ SEQUENCE 193 AA; 23001 MW; 3341AEF611A4478P CRC64;  
 -----  
 Query Match 76.5%; Score 13; DB 1; Length 193;  
 Best Local Similarity 50.0%; Pred. No. 4.1e+02;  
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 QY 2 EXXXLL 7  
 DB 33 EATLL 38  
 -----  
 RESULT 19  
 NDOC PSEPU  
 ID NDOC PSEPU STANDARD; PRT; 194 AA.  
 AC P23095; Q52125;  
 DT 01-NOV-1991 (Rel. 20, Created)  
 DT 01-NOV-1991 (Rel. 20, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Naphthalene 1,2-dioxigenase beta subunit (EC 1.14.12.12) (Naphthalene  
 1,2-dioxigenase ISP beta).  
 GN NDOC OR NAHAD OR DOXD.  
 OS Pseudomonas putida, and  
 OS Pseudomonas sp. (strain C18).  
 OG Plasmid pD7G1, and Plasmid NAH7.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
 OC Pseudomonadaceae; Pseudomonas.  
 OX NCBI\_TaxID=303, 306;  
 RN [1]  
 RC SEQUENCE FROM N.A.  
 RP SPECIES=P.Putida; STRAIN=NCIMB 9816;  
 RX MEDLINE=8211973; PubMed=3243438;  
 PA Kerkela S., Lehtvaeslahti H., Palva E.T., Teeri T.H.;  
 "Cloning, nucleotide sequence and characterization of genes encoding  
 naphthalene dioxigenase of Pseudomonas putida strain NCIB9816.";  
 [2]  
 RN Gene 73:355-362(1988).  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C18;  
 RX MEDLINE=94042852; PubMed=82266311;  
 RA Denome S.A., Stanley D.C., Olson E.S., Young K.D.;  
 "Metabolism of dibenzothiophene and naphthalene in Pseudomonas  
 strains: complete DNA sequence of an upper naphthalene catabolic  
 pathway.";  
 [3]  
 RN J. Bacteriol. 175:6890-6901(1993).  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=P.Putida; STRAIN=G7 / ATCC 17485; PLASMID=NAH7;  
 RX MEDLINE=93252277; PubMed=8486285;  
 RA Simon M.J., Oaslund T.D., Saunders R., Ensley B.D., Suggs S.,  
 Harcourt A., Suen W.-C., Cruden D.L., Gibson D.T., Zylstra G.J.;  
 "Sequences of genes encoding naphthalene dioxigenase in Pseudomonas  
 putida strains G7 and NCIB 9816-4.";  
 [4]  
 RN Gene 127:31-37(1993).  
 RP

-----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----  
 EMBL; M23914; AAB47592.1; -  
 EMBL; U49496; AAA92142.1; -  
 EMBL; M60405; AAA16127.1; -  
 EMBL; M83949; AAA25903.1; -  
 PIR; JN0645; JN0645.  
 PIR; JS0072; JS0072.  
 PDB; 1NDO; 23-MAR-99.  
 PDB; 1EG9; 31-MAY-00.  
 PDB; 1O7G; 20-FEB-03.  
 PDB; 1O7H; 20-FEB-03.  
 PDB; 1O7M; 20-FEB-03.  
 PDB; 1O7P; 20-FEB-03.  
 PDB; 1O7Q; 20-FEB-03.  
 InterPro: IPR00391; Ring\_hydroxyl\_B.  
 Pfam: PF00866; Ring\_hydroxyl\_B; 1.  
 ProDom: PD002199; Ring\_hydroxyl\_B; 1.  
 KW Aromatic hydrocarbons catabolism; Oxidoreductase; Dioxigenase; NAD;  
 Plasmid; 3D-structure.  
 FT VARIANT 17 17  
 FT VARIANT 19 19  
 FT VARIANT 28 28  
 FT VARIANT 39 40  
 FT VARIANT 99 99  
 FT VARIANT 109 109  
 FT VARIANT 127 127  
 FT VARIANT 130 130  
 FT VARIANT 132 132  
 FT TURN 132 132  
 FT TURN 135 135  
 FT TURN 137 137  
 FT HELIX 13 23  
 FT HELIX 28 46  
 FT TURN 47 48  
 FT HELIX 50 57  
 FT STRAND 58 69  
 -----  
 E -> Q (IN STRAIN G7).  
 I -> F (IN STRAIN G7).  
 S -> A (IN STRAIN G7).  
 TQ -> NR (IN STRAIN G7).  
 V -> I (IN STRAIN G7).  
 G -> S (IN STRAIN G7).  
 M -> R (IN STRAIN G7).  
 N -> D (IN STRAIN G7).  
 K -> E (IN STRAIN G7).  
 -----  
 [5]  
 RN X-RAY CRYSTALLOGRAPHY (2.25 ANGSTROMS).  
 RC SPECIES=P.Putida; STRAIN=NCIMB 9816;  
 RX MEDLINE=98298434; PubMed=9634695;  
 RA Kauppi B., Lee K., Carredano E., Parales R.E., Gibson D.T., Eklund H.,  
 Ramaswamy S.;  
 "Structure of an aromatic-ring-hydroxylating dioxigenase-naphthalene  
 1,2-dioxigenase.";  
 RL Structure 6:571-586(1998).  
 CC -|- FUNCTION: COMPONENT OF NAPHTHALENE DIOXYGENASE (NDO).  
 CC MULTICOMPONENT ENZYME SYSTEM WHICH CATALYZES THE INCORPORATION OF  
 CC BOTH ATOMS OF MOLECULAR OXYGEN INTO NAPHTHALENE TO FORM CIS-  
 CC NAPHTHALENE DIHYDRODIOL. THE BETA SUBUNIT MAY BE RESPONSIBLE FOR  
 CC THE SUBSTRATE SPECIFICITY OF THE ENZYME.  
 CC -|- CATALYTIC ACTIVITY: Naphthalene + NADH + O(2) -> (1R,2S)-1,2-  
 CC dihydronaphthalene-1,2-diol + NAD(+).  
 CC -|- COFACTOR: PROBABLY BINDS A 2FE-2S GROUP AND AN IRON ATOM.  
 CC -|- PATHWAY: UPPER NAPHTHALENE CATABOLIC PATHWAY WHICH INVOLVES  
 CC CONVERSION OF NAPHTHALENE TO SALICYLATE, AND CATABOLISM OF  
 CC DIBENZOTHIOPHENE (DBT) AND PHENANTHRENE. CONVERTS PHENANTHRENE TO  
 CC 1-HYDROXY-2-NAPHTHOIC ACID AND THE METABOLISM OF DBT IS LIMITED TO  
 CC OXIDATION OF THE AROMATIC RING.  
 CC -|- SUBUNIT: NAPHTHALENE DIOXYGENASE (NDO) MULTICOMPONENT ENZYME  
 CC SYSTEM IS COMPOSED OF AN ELECTRON TRANSFER COMPONENT AND AN IRON  
 CC SULFUR PROTEIN (ISP). THE ELECTRON TRANSFER COMPONENT IS COMPOSED  
 CC OF FERREDOXIN REDUCTASE (NDR) AND FERREDOXIN (NDOA), AND ISP IS  
 CC COMPOSED OF AN HEXAMER OF THREE LARGE ALPHA SUBUNITS (NDOB) AND  
 CC THREE SMALL BETA SUBUNITS (NDOC).  
 CC -|- SIMILARITY: BELONGS TO THE BACTERIAL RING-HYDROXYLATING  
 CC DIOXYGENASE BETA SUBUNIT FAMILY.  
 -----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----  
 EMBL; M23914; AAB47592.1; -  
 EMBL; U49496; AAA92142.1; -  
 EMBL; M60405; AAA16127.1; -  
 EMBL; M83949; AAA25903.1; -  
 PIR; JN0645; JN0645.  
 PIR; JS0072; JS0072.  
 PDB; 1NDO; 23-MAR-99.  
 PDB; 1EG9; 31-MAY-00.  
 PDB; 1O7G; 20-FEB-03.  
 PDB; 1O7H; 20-FEB-03.  
 PDB; 1O7M; 20-FEB-03.  
 PDB; 1O7P; 20-FEB-03.  
 PDB; 1O7Q; 20-FEB-03.  
 InterPro: IPR00391; Ring\_hydroxyl\_B.  
 Pfam: PF00866; Ring\_hydroxyl\_B; 1.  
 ProDom: PD002199; Ring\_hydroxyl\_B; 1.  
 KW Aromatic hydrocarbons catabolism; Oxidoreductase; Dioxigenase; NAD;  
 Plasmid; 3D-structure.  
 FT VARIANT 17 17  
 FT VARIANT 19 19  
 FT VARIANT 28 28  
 FT VARIANT 39 40  
 FT VARIANT 99 99  
 FT VARIANT 109 109  
 FT VARIANT 127 127  
 FT VARIANT 130 130  
 FT VARIANT 132 132  
 FT TURN 132 132  
 FT TURN 135 135  
 FT TURN 137 137  
 FT HELIX 13 23  
 FT HELIX 28 46  
 FT TURN 47 48  
 FT HELIX 50 57  
 FT STRAND 58 69  
 -----  
 E -> Q (IN STRAIN G7).  
 I -> F (IN STRAIN G7).  
 S -> A (IN STRAIN G7).  
 TQ -> NR (IN STRAIN G7).  
 V -> I (IN STRAIN G7).  
 G -> S (IN STRAIN G7).  
 M -> R (IN STRAIN G7).  
 N -> D (IN STRAIN G7).  
 K -> E (IN STRAIN G7).

FT TURN 74 75  
 FT STRAND 84 90  
 FT HELIX 92 103  
 FT TURN 105 106  
 FT HELIX 108 110  
 FT STRAND 115 127  
 FT TURN 132 133  
 FT STRAND 134 147  
 FT TURN 148 150  
 FT STRAND 151 165  
 FT TURN 167 168  
 FT STRAND 171 179  
 FT STRAND 185 185  
 SQ SEQUENCE 194 AA; 22935 MW; 168D2E1535E66416 CRC64;

Query Match 76.5%; Score 13; DB 1; Length 194;  
 Best Local Similarity 50.0%; Pred. No. 4.1e+02;  
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

2 EXXXLL 7  
 33 EATLL 38

## RESULT 20

PF03 ARATH STANDARD; PRT; 195 AA.  
 AC P57741;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Probable prefoldin subunit 3.  
 GN AT5G49510 OR K6M13.5.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 OX NCBI\_TaxID=3702;  
 EN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia;  
 RA MEDLINE=20181125; PubMed=10718197;  
 RX Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Kotani H.,  
 RA Tabata S.  
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. X. Sequence  
 RT features of the regions of 3,076,755 bp covered by sixty P1 and TAC  
 RT clones.";  
 RI DNA Res. 7:31-63(2000).

-1- FUNCTION: BINDS SPECIFICALLY TO CYTOSOLIC CHAPERONIN (C-CPN) AND  
 TRANSFERS TARGET PROTEINS TO IT. BINDS TO NASCENT POLYPEPTIDE  
 CHAIN AND PROMOTES FOLDING IN AN ENVIRONMENT IN WHICH THERE ARE  
 MANY COMPETING PATHWAYS FOR NONNATIVE PROTEINS (BY SIMILARITY).  
 -1- SUBUNIT: HETEROHEXAMER OF TWO PFD-ALPHA TYPE AND FOUR PFD-BETA  
 TYPE SUBUNITS (BY SIMILARITY).

-1- SIMILARITY: BELONGS TO THE PREPOLDIN ALPHA SUBUNIT FAMILY.

-----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; AB023033; BAB10764.1;  
 InterPro; IPR004127; DUF232.  
 Pfam; PF02996; Prefoldin; 1.  
 KW Chaperone.

SQ SEQUENCE 195 AA; 21465 MW; C1F018964AA6758E CRC64;

Query Match 76.5%; Score 13; DB 1; Length 195;  
 Best Local Similarity 50.0%; Pred. No. 4.1e+02;  
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7  
 Db 132 EASALL 137

## RESULT 21

FIMW SALTY STANDARD; PRT; 198 AA.  
 ID FIMW SALTY  
 AC P37928;  
 DT 01-OCT-1994 (Rel. 30, Created)  
 DT 01-OCT-1994 (Rel. 30, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Fimbriae W protein.  
 GN FIMW OR STM552.  
 OS Salmonella typhimurium.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; Salmonella.  
 OX NCBI\_TaxID=602;  
 EN [1]  
 RP SEQUENCE FROM N.A.  
 RC Swenson D.L., Clegg S.;  
 RA Submitted (JUN-1993) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=LT2 / SGSC1412 / ATCC 700720;  
 RX MEDLINE=21534948; PubMed=11677609;  
 RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,  
 RA Courtney L., Porwollik S., Ali J., Dante M., Du P., Hou S., Layman D.,  
 RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,  
 RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,  
 RA Waterston R., Wilson R.K.;  
 RT "Complete genome sequence of Salmonella enterica serovar Typhimurium  
 RT LT2.";  
 RL Nature 413:852-856(2001).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; L19338; AAA75424.1;  
 DR EMBL; AB008721; AAL19506.1;  
 DR StyGene; SG10509; fimm.  
 DR InterPro; IPR000792; HTH LuxR.  
 DR SMART; SM00421; HTH\_LUXR; 1.  
 KW Fimbria; Complete proteome.

SQ SEQUENCE 198 AA; 23436 MW; 46DC977A312CE292 CRC64;

Query Match 76.5%; Score 13; DB 1; Length 198;  
 Best Local Similarity 50.0%; Pred. No. 4.2e+02;  
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7  
 Db 153 ETPALL 158

## RESULT 22

EPA2 CHICK STANDARD; PRT; 200 AA.  
 ID EPA2 CHICK  
 AC P52802;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE Ephrin-A2 precursor (EPH-related receptor tyrosine kinase ligand 6)  
 DE (LERK-6) (ELF-1)  
 GN EFNA2 OR EPLA6 OR LERK6 OR ELF1.  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Archosauzia; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Gallus.  
 OX NCBI\_TaxID=9031;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=95360981; PubMed=7634327;  
 RA Cheng H.J., Nakamoto M., Bergemann A.D., Planagan J.G.,  
 RT "Complementary gradients in expression and binding of ELP-1 and Mek4  
 in development of the topographic retinorectal projection map.";  
 RL Cell 82:371-381(1995).  
 CC  
 CC -1- SUBUNIT: BINDS TO THE RECEPTOR TYROSINE KINASES EPHA3, EPHA4 AND  
 EPHAS (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor  
 (Potential).  
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN A GRADIENT ACROSS THE TECTUM  
 BEING MORE STRONGLY EXPRESSED AT THE POSTERIOR POLE.  
 CC -1- SIMILARITY: BELONGS TO THE EPHRIN FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: L40932; AAC42229.1; .  
 DR InterPro; IPR001799; Ephrin.  
 DR Pfam; PF00812; Ephrin; 1.  
 DR PRINTS; PR01347; EPHRIN.  
 DR ProDom; PD002533; Ephrin; 1.  
 DR PROSITE; PS01299; EPHRIN; 1.  
 KW Glycoprotein; GPI-anchor; Signal.  
 PT SIGNAL 1 22 POTENTIAL.  
 FT CHAIN 23 200 EPHRIN-A2.  
 FT CARBOHYD 36 36 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 161 161 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 175 175 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 200 AA; 23049 MW; 8FAB1A5E45B9D96 CRC64;  
 Query Match 76.5%; Score 13; DB 1; Length 200;  
 Best Local Similarity 50.0%; Pred. No. 4.2e+02;  
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 OY 2 EXXXLL 7  
 DB 5 EAAALL 10  
 RESULT 23  
 ID YL96 RALSO STANDARD; PRT; 207 AA.  
 AC P58634;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Maf-like protein RSC2196.  
 GN RSC2196 OR RS01403.  
 OS Ralstonia solanacearum (Pseudomonas solanacearum).  
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;  
 OC Burkholderiaceae; Ralstonia.  
 OX NCBI\_TaxID=305;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=GM1000;  
 RX MEDLINE=21681879; PubMed=11823852;  
 RA Salanoubat M., Genin S., Attiguenave F., Gouzy J., Mangenot S.,  
 RA Arlat M., Billault A., Brotier P., Camus J.C., Cattolico L.,  
 RA Chandler M., Choisme N., Claudel-Renard C., Cunnac S., Demange N.,  
 RA Gapin C., Lavie M., Moisan A., Robert C., Saurin W., Schiex T.,  
 RA Siguer P., Thebaud P., Whalen M., Wincker P., Levy M.,  
 RA Weissenbach J., Boucher C.A.;  
 RT "Genome sequence of the plant pathogen Ralstonia solanacearum.";

RL Nature 415:497-502(2002).  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).  
 CC -1- SIMILARITY: BELONGS TO THE MAP FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: AL646068; CAD15903.1; .  
 DR HAMAP; MF\_00528; .; 1.  
 DR InterPro; IPR003697; Maf.  
 DR Pfam; PF02545; Maf; 1.  
 DR TIGRPFAMs; TIGR00172; maf; 1.  
 KW Complete proteome.  
 FT ACT\_SITE 46 46 POTENTIAL.  
 SQ SEQUENCE 207 AA; 21866 MW; 0503C75A1522CA55 CRC64;  
 Query Match 76.5%; Score 13; DB 1; Length 207;  
 Best Local Similarity 50.0%; Pred. No. 4.4e+02;  
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 OY 2 EXXXLL 7  
 DB 195 ETAALL 200  
 RESULT 24  
 ID VATE CHLPN STANDARD; PRT; 208 AA.  
 AC Q9Z995;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE V-type ATP synthase subunit E (EC 3.6.3.14) (V-type ATPase subunit E).  
 GN ATP6 OR CPN0086 OR CP0688.  
 OS Chlamydia pneumoniae (Chlamydophila pneumoniae).  
 OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydophila.  
 OX NCBI\_TaxID=83558;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CWL029;  
 RX MEDLINE=99206606; PubMed=10192388;  
 RA Kalman S., Mitchell W., Marathe R., Lammel C., Fan J., Hyman R.W.,  
 RA Olinger L., Grimwood J., Davis R.W., Stephens R.S.;  
 RT "Comparative genomes of Chlamydia pneumoniae and C. trachomatis.";  
 RL Nat. Genet. 21:385-389(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=AB39;  
 RX MEDLINE=20150255; PubMed=10684935;  
 RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,  
 RA White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,  
 RA Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,  
 RA Gwinn M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,  
 RA Eisen J., Fraser C.M.;  
 RT "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia  
 pneumoniae AR39.";  
 RL Nucleic Acids Res. 28:1397-1406(2000).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=J138;  
 RX MEDLINE=20330349; PubMed=10871362;  
 RA Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,  
 RA Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;  
 RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138  
 from Japan and CWL029 from USA.";  
 RL Nucleic Acids Res. 28:2311-2314(2000).  
 CC -1- FUNCTION: PRODUCES ATP FROM ADP IN THE PRESENCE OF A PROTON  
 GRADIENT ACROSS THE MEMBRANE.

```

CC -1- CATALYTIC ACTIVITY: ATP + H(2)O + H(+) (In) = ADP + phosphate +
CC H(+) (Out).
CC -1- SIMILARITY: Belongs to the V-ATPase E subunit family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE001594; RAD18239.1; -
CC EMBL; AE002227; RAF73697.1; -
CC EMBL; AP002545; EAA98296.1; -
CC PIR; F6501; F66501.
CC TIGR; CP0688; -.
CC HAMAP; MF_00311; -.
CC InterPro; IPR002842; ATPsynth_Eub.
CC Pfam; PF01991; V-ATP-synt_E_1.
CC Hydrolase; ATP synthase; Hydrogen ion transport; Complete proteome.
CC SEQUENCE 208 AA; 23053 MW; FC93ECC17DE14903 CRC64;
CC -----
CC Query Match 76.5%; Score 13; DB 1; Length 208;
CC Best Local Similarity 50.0%; Pred.No. 4.4e+02;
CC Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
CC -----
CC QY 2 EXXXLL 7
CC | | |
CC Db 28 EAAALL 33
CC -----
CC RESULT 25
CC ID YFDX_ECOLI STANDARD; PRT; 211 AA.
CC AC P58403;
CC DT 28-FEB-2003 (Rel. 41, Created)
CC DT 28-FEB-2003 (Rel. 41, Last sequence update)
CC DT 28-FEB-2003 (Rel. 41, Last annotation update)
CC DE YfjX protein precursor.
CC GN YFDX OR Z3640 OR ECS3255.
CC OS Escherichia coli O157:H7.
CC OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
CC OC Enterobacteriaceae; Escherichia.
CC OX NCBI_TaxID=83334;
CC [1]
CC [2]
CC SEQUENCE FROM N.A.
CC STRAIN=O157:H7 / EDL933 / ATCC 700927;
CC MEDLINE=21074935; PubMed=11206551;
CC Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
CC Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
CC Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
CC Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamoudis K.,
CC Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
CC Welch R.A., Blattner F.R.;
CC "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7."
CC Nature 409:529-533(2001).
CC [2]
CC SEQUENCE FROM N.A.
CC STRAIN=O157:H7 / RIMD 0509952;
CC MEDLINE=21156231; PubMed=11258796;
CC Hayaishi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
CC Hayashi T., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
CC Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
CC Kuhara S., Shiba T., Hattori M., Shinagawa H.;
CC "Complete genome sequence of enterohaemorrhagic Escherichia coli
CC O157:H7 and genomic comparison with a laboratory strain K-12."
CC DNA Res. 8:11-22(2001).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```

```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE005469; AAG57501.1; -
CC EMBL; AP002561; BAB36678.1; -
CC PIR; A85880; A85880.
CC PIR; G91035; G91035.
CC Signal; Complete proteome.
CC SIGNAL 1 21
CC CHAIN 22 211 YFDX PROTEIN.
CC SEQUENCE 211 AA; 22963 MW; E56971BF98B7DEE1 CRC64;
CC -----
CC Query Match 76.5%; Score 13; DB 1; Length 211;
CC Best Local Similarity 50.0%; Pred.No. 4.5e+02;
CC Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
CC -----
CC QY 2 EXXXLL 7
CC | | |
CC Db 81 EASALL 86
CC -----
CC RESULT 26
CC YFDX_ECOLI
CC ID YFDX_ECOLI STANDARD; PRT; 211 AA.
CC AC P76570;
CC DT 16-OCT-2001 (Rel. 40, Created)
CC DT 16-OCT-2001 (Rel. 40, Last sequence update)
CC DT 28-FEB-2003 (Rel. 41, Last annotation update)
CC DE YfjX protein precursor.
CC GN YFDX OR B2375.
CC OS Escherichia coli.
CC OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
CC OC Enterobacteriaceae; Escherichia.
CC OX NCBI_TaxID=562;
CC [1]
CC [2]
CC SEQUENCE FROM N.A.
CC STRAIN=K12 / MG1655;
CC MEDLINE=97426617; PubMed=9278503;
CC Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
CC Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
CC Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
CC Mau B., Shao Y.;
CC "The complete genome sequence of Escherichia coli K-12."
CC Science 277:1453-1474(1997).
CC [2]
CC SEQUENCE OF 22-31.
CC STRAIN=K12 / W3104;
CC MEDLINE=21101830; PubMed=11157960;
CC Nishino K., Yamaguchi A.;
CC "Overexpression of the response regulator evgA of the two-component
CC signal transduction system modulates multidrug resistance conferred
CC by multidrug resistance transporters."
CC J. Bacteriol. 183:1455-1458(2001).
CC -1- INDUCTION: By evgA.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE000326; AAC75434.1; -
CC PIR; D65011; D65011.
CC EcoGene; EGI4146; yfdX.
CC Signal; Complete proteome.
CC SIGNAL 1 21
CC CHAIN 22 211 YFDX PROTEIN.
CC SEQUENCE 211 AA; 22977 MW; D2750C9B578B2D2E CRC64;

```

Query Match 76.5%; Score 13; DB 1; Length 211;  
 Best Local Similarity 50.0%; Pred. No. 4.8e+02;  
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7  
 DB 81 EASALL 86

## RESULT 27

ATP6\_ANAPL STANDARD; PRT; 227 AA.  
 AC PS0654;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DS ATP synthase A chain (EC 3.6.3.14) (Protein 6).  
 GN MTATP6 OR ATP6.  
 OS Anas platyrhynchos (Domestic duck).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archaeosuria; Aves; Neognathae; Anseriformes; Anatidae; Anas.  
 OX NCBI\_TaxID=8839;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Pekin breed; TISSUE=Liver;  
 RX MEDLINE=94047124; PubMed=8230253;  
 RA Ramirez V., Savio P., Morais R.;  
 RT "Molecular characterization and evolution of a duck mitochondrial  
 RT genome.";  
 RL J. Mol. Evol. 37:296-310(1993).  
 CC -1- FUNCTION: KEY COMPONENT OF THE PROTON CHANNEL; IT MAY PLAY A  
 CC DIRECT ROLE IN THE TRANSLLOCATION OF PROTONS ACROSS THE MEMBRANE.  
 CC -1- CATALYTIC ACTIVITY: ATP + H(2)O + H(+) (In) = ADP + phosphate +  
 CC H(+) (Out).  
 CC -1- SUBUNIT: F-TYPE ATPASES HAVE 2 COMPONENTS, CF(1) - THE CATALYTIC  
 CC CORE - AND CF(0) - THE MEMBRANE PROTON CHANNEL. CF(1) HAS FIVE  
 CC SUBUNITS: ALPHA(3), BETA(3), GAMMA(1), DELTA(1), EPSILON(1). CF(0)  
 CC HAS THREE MAIN SUBUNITS: A, B AND C.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -1- SIMILARITY: BELONGS TO THE ATPASE A CHAIN FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----  
 DR EMBL; L22476; AAA72036.1; --  
 DR HSP; P00855; ICI7.  
 DR InterPro; IPR000568; ATPsynth\_Asub.  
 DR Pfam; PF00119; ATP-synt\_A; 1.  
 DR PRINTS; PR00123; ATPASEA.  
 DR TIGRFAMs; TIGR01131; ATP synt 6 or A; 1.  
 DR PROSITE; PS00449; ATPASE\_A; 1\_--  
 DR PROSITE; PS00449; ATPASE\_A; 1\_--  
 DR Hydrogen ion transport; CF(0); Mitochondrion; Transmembrane.  
 SQ SEQUENCE 227 AA; 24997 MW; DD136234EB2E87FB CRC64;

Query Match 76.5%; Score 13; DB 1; Length 227;  
 Best Local Similarity 50.0%; Pred. No. 4.8e+02;  
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7  
 DB 146 ETTSLL 151

## RESULT 28

ATP6\_CHICK STANDARD; PRT; 227 AA.  
 AC F14092;

DT 01-JAN-1990 (Rel. 13, Created)  
 DT 01-JAN-1990 (Rel. 13, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE ATP synthase A chain (EC 3.6.3.14) (Protein 6).  
 GN MTATP6 OR ATP6.  
 OS Gallus gallus (Chicken).  
 OC Mitochondrion.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archaeosuria; Aves; Neognathae; Galliformes; Phasianinae;  
 OC Gallus.  
 OX NCBI\_TaxID=9031;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=89342447; PubMed=2474659;  
 RA Desjardins P., L'Abbe D., Lang B.F., Morais R.;  
 RT "Putative chicken 'muscle-specific 7 S RNA' is related to the  
 RT mitochondrial ATPase 6 gene.";  
 RL J. Mol. Biol. 207:625-629(1989).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=90230301; PubMed=2329578;  
 RA Desjardins P., Morais R.;  
 RT "Sequence and gene organization of the chicken mitochondrial genome.  
 RT A novel gene order in higher vertebrates.";  
 RL J. Mol. Biol. 212:599-634(1990).  
 CC -1- FUNCTION: KEY COMPONENT OF THE PROTON CHANNEL; IT MAY PLAY A  
 CC DIRECT ROLE IN THE TRANSLLOCATION OF PROTONS ACROSS THE MEMBRANE.  
 CC -1- CATALYTIC ACTIVITY: ATP + H(2)O + H(+) (In) = ADP + phosphate +  
 CC H(+) (Out).  
 CC -1- SUBUNIT: F-TYPE ATPASES HAVE 2 COMPONENTS, CF(1) - THE CATALYTIC  
 CC CORE - AND CF(0) - THE MEMBRANE PROTON CHANNEL. CF(1) HAS FIVE  
 CC SUBUNITS: ALPHA(3), BETA(3), GAMMA(1), DELTA(1), EPSILON(1). CF(0)  
 CC HAS THREE MAIN SUBUNITS: A, B AND C.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -1- SIMILARITY: BELONGS TO THE ATPASE A CHAIN FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----  
 DR EMBL; X15841; CAA33842.1; --  
 DR EMBL; X52392; CAA36630.1; --  
 DR PIR; S10192; S10192.  
 DR HSP; P00855; ICI7.  
 DR InterPro; IPR000568; ATPsynth\_Asub.  
 DR Pfam; PF00119; ATP-synt\_A; 1.  
 DR PRINTS; PR00123; ATPASEA.  
 DR TIGRFAMs; TIGR01131; ATP synt 6 or A; 1.  
 DR PROSITE; PS00449; ATPASE\_A; 1\_--  
 DR Hydrogen ion transport; CF(0); Mitochondrion; Transmembrane.  
 SQ SEQUENCE 227 AA; 24797 MW; E3BCSE98A84385D CRC64;

Query Match 76.5%; Score 13; DB 1; Length 227;  
 Best Local Similarity 50.0%; Pred. No. 4.8e+02;  
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7  
 DB 146 ETTSLL 151

RESULT 29  
 ATP6\_STRCA STANDARD; PRT; 227 AA.  
 AC O21402;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE ATP synthase A chain (EC 3.6.3.14) (Protein 6).  
 GN MTATP6 OR ATP6.



```

ID  NAGB LISMO          STANDARD;          PRT;   234 AA.
AC  Q8Y8E7;
DT  15-SEP-2003 (Rel. 42, Created)
DT  15-SEP-2003 (Rel. 42, Last sequence update)
DT  15-SEP-2003 (Rel. 42, Last annotation update)
DE  Glucosamine-6-phosphate deaminase (EC 3.5.99.6) (Glucosamine-6-
GN  phosphate isomerase) (GNPDA) (GlcN6P deaminase).
OS  NAGB OR LMO0957.
OC  Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
OX  NCBI_TaxID=1639;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=EGD-e / Serovar 1/2a;
RX  MEDLINE=21517279; PubMed11679669;
RA  Glaser P., Frangeul L., Buchrieser C., Rueniok C., Amend A.,
RA  Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
RA  Charbit A., Chatouani F., Couve E., de Daruvar A., Dehoux P.,
RA  Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dusauguet O.,
RA  Gautier K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,
RA  Gauthier L., Goebel N., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA  Jones L.-M., Kaerst U., Kraft J., Kuhn M., Kunst F., Kurapkat G.,
RA  Madueno E., Maitoum A., Mata Vicente J., Ng E., Medjari H.,
RA  Nordiek G., Novella S., de Pablo B., Perez-Diaz J.-C., Purcell R.,
RA  Remmel B., Rose M., Schluter T., Simoes N., Tierrez A.,
RA  Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.,
RA  "Comparative genomics of Listeria species.";
RL  Science 294:849-852(2001).
CC  -!- FUNCTION: Catalyzes the reversible isomerization-deamination of
CC  glucosamine 6-phosphate (GlcN6P) to form fructose 6-phosphate
CC  (Fru6P) and ammonium ion (By similarity).
CC  -!- CATALYTIC ACTIVITY: D-glucosamine 6-phosphate + H(2)O = D-fructose
CC  6-phosphate + NH(3).
CC  -!- PATHWAY: N-acetylglucosamine utilization.
CC  -!- SIMILARITY: Belongs to the glucosamine/galactosamine-6-phosphate
CC  isomerase family. NagB subfamily.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; AL591977; CAC99035.1; --
DR  ListList; LMO00957; --
DR  HAMAP; MF_01241; --; 1.
DR  InterPro; IPR006148; Gluc_gal_isom.
DR  Pfam; PF01182; Glucosamine iso; 1.
DR  TIGRfam; TIGR00502; nagB_1.
DR  PROSITE; PS01161; GLC_GALNAC_ISOMERASE; FALSE NEG.
KW  Carbohydrate metabolism; Hydrolyase; Complete proteome.
FT  ACT_SITE 63 63
FT  ENOLIZATION STEP (BY SIMILARITY).
FT  PART OF THE CATALYTIC TRIAD (BY
FT  SIMILARITY).
FT  GENERAL BASE INVOLVED IN THE CATALYSIS OF
FT  THE RING-OPENING STEP OF GLCN6P; PART OF
FT  THE CATALYTIC TRIAD (BY SIMILARITY).
FT  PART OF THE CATALYTIC TRIAD (BY
FT  SIMILARITY).
FT  ACT_SITE 129 129
FT  ACT_SITE 131 131
FT  ACT_SITE 136 136
FT  SEQUENCE 234 AA; 25510 MW; 4E274CC86378160 CRC64;
Query Match 76.5%; Score 13; DB 1; Length 234;
Best Local Similarity 50.0%; Pred. No. 4.9e+02;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 2 EXXXLL 7
DB 229 EAASLL 234

```

```

RESULT 32
SFSA_VIBPA
ID  SFSA_VIBPA          STANDARD;          PRT;   238 AA.
AC  Q87LV9;
DT  15-SEP-2003 (Rel. 42, Created)
DT  15-SEP-2003 (Rel. 42, Last sequence update)
DT  15-SEP-2003 (Rel. 42, Last annotation update)
DE  Sugar fermentation stimulation protein homolog.
GN  SFSA OR VP2499.
OS  Vibrio parahaemolyticus.
OC  Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC  Vibrionaceae; Vibrio.
OX  NCBI_TaxID=670;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=RIMD 2210633 / Serotype O3:K6;
RX  MEDLINE=22508454; PubMed12620739;
RA  Makino K., Oshima K., Kurokawa K., Yokoyama K., Uda T., Tagomori K.,
RA  Iijima Y., Najiima M., Nakano M., Yamashita A., Kubota Y., Kimura S.,
RA  Yasunaga T., Honda T., Shinagawa H., Hattori M., Iida T.;
RT  "Genome sequence of Vibrio parahaemolyticus: a pathogenic mechanism
RT  distinct from that of V. cholerae.";
RL  Lancet 361:743-749(2003).
CC  -!- SIMILARITY: BELONGS TO THE SFSA FAMILY.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; AP005081; BAC60762.1; --
DR  HAMAP; MF_00095; --; 1.
KW  Complete proteome.
SQ  SEQUENCE 238 AA; 26409 MW; 832B12CFC71A777 CRC64;
Query Match 76.5%; Score 13; DB 1; Length 238;
Best Local Similarity 50.0%; Pred. No. 5e+02;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 EXXXLL 7
DB 8 ESATLL 13

```

```

RESULT 33
VMT1_INBAC
ID  VMT1_INBAC          STANDARD;          PRT;   248 AA.
AC  P13879;
DT  01-JAN-1990 (Rel. 13, Created)
DT  01-JAN-1990 (Rel. 13, Last sequence update)
DT  01-MAY-1992 (Rel. 22, Last annotation update)
DE  Matrix protein M1.
OC  Influenza B virus (strain B/Ann Arbor/1/66 [cold-adapted]).
OC  Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC  Influenzavirus B.
OX  NCBI_TaxID=11522;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=88179548; PubMed3354202;
RA  Deborde D.C., Donabedian A.M., Herlocher M.L., Naeve C.W.,
RA  Maassab H.F.;
RT  "Sequence comparison of wild-type and cold-adapted B/Ann Arbor/1/66
RT  Influenza virus genes";
RL  Virology 163:429-443(1988).
CC  -!- FUNCTION: THIS PROTEIN FORMS A CONTINUOUS SHELL ON THE INNER SIDE
CC  OF THE LIPID BILAYER, BUT ITS FUNCTION IS UNCLEAR.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -

```



CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 CC ENBL; M20175; AAA66414.1; -  
 CC DR PIR; A30064; MFIVBC.  
 CC DR HSSP; P03485; 1AA7.  
 CC DR InterPro; IPR001561; Flu\_M1.  
 CC DR Pfam; PF00598; Flu\_M1; 1.  
 CC DR ProDom; PD001061; Flu\_M1; 1.  
 CC KW Matrix protein.  
 CC SQ SEQUENCE 248 AA; 27374 MW; 7143EP8FB45A594A CRC64;

Query Match 76.5%; Score 13; DB 1; Length 248;  
 Best Local Similarity 50.0%; Pred. No. 5.2e+02;  
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 2 EXXXLL 7  
 120 ESSALL 125

## RESULT 34

CC VMT1\_INBAD STANDARD; PRT; 248 AA.  
 CC AC P13880;  
 CC DT 01-JAN-1990 (Rel. 13, Created)  
 CC DT 01-JAN-1990 (Rel. 13, Last sequence update)  
 CC DT 01-MAY-1992 (Rel. 22, Last annotation update)  
 CC DE Matrix protein M1.

CC OS Influenza B virus (strain B/Ann Arbor/1/66 [wild-type]).  
 CC OC Viruses; sRNA negative-strand viruses; Orthomyxoviridae;  
 CC OC Influenzavirus B.  
 CC OX NCBI\_TaxID=11523;  
 CC RN [1]

CC SEQUENCE FROM N.A.

CC RX MEDLINE=88179548; PubMed=3354202;  
 CC RA Deborde D.C., Donabedian A.M., Herlocher M.L., Naeve C.W.,  
 CC RA Maassab H.F.;

CC RT Influenza virus Genes";  
 CC RT Virology 163:429-443(1988).  
 CC CC -1- FUNCTION: THIS PROTEIN FORMS A CONTINUOUS SHELL ON THE INNER SIDE.  
 CC OF THE LIPID BILAYER, BUT ITS FUNCTION IS UNCLEAR.

CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 CC ENBL; M20176; AAA66416.1; -  
 CC DR PIR; C30064; MFIVBW.  
 CC DR HSSP; P03485; 1AA7.  
 CC DR InterPro; IPR001561; Flu\_M1.  
 CC DR Pfam; PF00598; Flu\_M1; 1.  
 CC DR ProDom; PD001061; Flu\_M1; 1.  
 CC KW Matrix protein.  
 CC SQ SEQUENCE 248 AA; 27415 MW; 714F22C7319711C3 CRC64;

Query Match 76.5%; Score 13; DB 1; Length 248;  
 Best Local Similarity 50.0%; Pred. No. 5.2e+02;  
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 2 EXXXLL 7  
 120 ESSALL 125

CC VMT1\_INBAD STANDARD; PRT; 248 AA.  
 CC AC P06816;  
 CC DT 01-JAN-1988 (Rel. 06, Created)  
 CC DT 01-JAN-1988 (Rel. 06, Last sequence update)  
 CC DT 01-MAY-1992 (Rel. 22, Last annotation update)  
 CC DE Matrix protein M1.

CC OS Influenza B virus (strain B/Singapore/222/79).  
 CC OC Viruses; sRNA negative-strand viruses; Orthomyxoviridae;  
 CC OC Influenzavirus B.  
 CC OX NCBI\_TaxID=11544;  
 CC RN [1]

CC SEQUENCE FROM N.A.

CC RX MEDLINE=87071690; PubMed=3788064;  
 CC RA Hiebert S.W., Williams M.A., Lamb R.A.;

CC RT Nucleotide sequence of RNA segment 7 of influenza  
 CC B/Singapore/222/79; maintenance of a second large open reading  
 CC frame.";  
 CC RT Virology 155:747-751(1986).  
 CC CC -1- FUNCTION: THIS PROTEIN FORMS A CONTINUOUS SHELL ON THE INNER SIDE  
 CC OF THE LIPID BILAYER, BUT ITS FUNCTION IS UNCLEAR.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

## RESULT 35

CC VMT1\_INBLE STANDARD; PRT; 248 AA.  
 CC AC P03489;  
 CC DT 21-JUL-1986 (Rel. 01, Created)  
 CC DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 CC DT 01-MAY-1992 (Rel. 22, Last annotation update)  
 CC DE Matrix protein M1.  
 CC OS Influenza B virus (strain B/Lee/40).  
 CC OC Viruses; sRNA negative-strand viruses; Orthomyxoviridae;  
 CC OC Influenzavirus B.  
 CC OX NCBI\_TaxID=11535;  
 CC RN [1]

CC SEQUENCE FROM N.A.

CC RX MEDLINE=82154757; PubMed=6278729;  
 CC RA Briedis D.J., Lamb R.A., Choppin P.W.;

CC RT "Sequence of RNA segment 7 of the influenza B virus genome: partial  
 CC amino acid homology between the membrane proteins (M1) of influenza A  
 CC and B viruses and conservation of a second open reading frame.";  
 CC RT Virology 116:581-588(1982).  
 CC CC -1- FUNCTION: THIS PROTEIN FORMS A CONTINUOUS SHELL ON THE INNER SIDE  
 CC OF THE LIPID BILAYER, BUT ITS FUNCTION IS UNCLEAR.

CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC ENBL; J02094; AAA43726.1; -  
 CC DR PIR; A04082; MFIVL.  
 CC DR HSSP; P03485; 1AA7.  
 CC DR InterPro; IPR001561; Flu\_M1.  
 CC DR Pfam; PF00598; Flu\_M1; 1.  
 CC DR ProDom; PD001061; Flu\_M1; 1.  
 CC KW Matrix protein.  
 CC SQ SEQUENCE 248 AA; 27554 MW; 46076P9ACBFB73CB CRC64;

Query Match 76.5%; Score 13; DB 1; Length 248;  
 Best Local Similarity 50.0%; Pred. No. 5.2e+02;  
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 2 EXXXLL 7  
 120 ESSALL 125

## RESULT 36

CC VMT1\_INBSI STANDARD; PRT; 248 AA.  
 CC AC P06816;  
 CC DT 01-JAN-1988 (Rel. 06, Created)  
 CC DT 01-JAN-1988 (Rel. 06, Last sequence update)  
 CC DT 01-MAY-1992 (Rel. 22, Last annotation update)  
 CC DE Matrix protein M1.

CC OS Influenza B virus (strain B/Singapore/222/79).  
 CC OC Viruses; sRNA negative-strand viruses; Orthomyxoviridae;  
 CC OC Influenzavirus B.  
 CC OX NCBI\_TaxID=11544;  
 CC RN [1]

CC SEQUENCE FROM N.A.

CC RX MEDLINE=87071690; PubMed=3788064;  
 CC RA Hiebert S.W., Williams M.A., Lamb R.A.;

CC RT Nucleotide sequence of RNA segment 7 of influenza  
 CC B/Singapore/222/79; maintenance of a second large open reading  
 CC frame.";  
 CC RT Virology 155:747-751(1986).  
 CC CC -1- FUNCTION: THIS PROTEIN FORMS A CONTINUOUS SHELL ON THE INNER SIDE  
 CC OF THE LIPID BILAYER, BUT ITS FUNCTION IS UNCLEAR.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; M14909; AAA67100.1; -  
 DR PIR; A25619; MP1VBI.  
 DR HSP; P03485; IAA7.  
 DR InterPro; IPR001561; Flu\_M1.  
 DR Pfam; PF00598; Flu\_M1; 1.  
 DR ProDom; PD001061; Flu\_M1; 1.  
 KW Matrix protein.  
 SQ SEQUENCE 248 AA; 27474 MW; 514EACFC5C21CAAD CRC64;

Query Match 76.5%; Score 13; DB 1; Length 248;  
 Best Local Similarity 50.0%; Pred. No. 5.2e+02;  
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

2 EXXXLL 7

Db 120 ESSALL 125

# RESULT 37

ID UDP\_ECOLI

AC P12758; STANDARD; PRT; 252 AA.

DT 01-OCT-1989 (Rel. 12, Created)

DT 01-OCT-1993 (Rel. 27, Last sequence update)

DT 15-SEP-2003 (Rel. 42, Last annotation update)

DE Uridine phosphorylase (EC 2.4.2.3) (UDPase).

GN UDP OR B3831

OS Escherichia coli.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

OC Enterobacteriaceae; Escherichia.

OX NCBI\_TaxID=562;

RN [1] SEQUENCE FROM N.A.

RP STRAIN=K12;

RX MEDLINE=89386014; PubMed=2674901;

RA Walton L., Richards C.A., Elwell L.P.;

RT "Nucleotide sequence of the Escherichia coli uridine phosphorylase

RL (udp) gene.";

RL Nucleic Acids Res. 17:6741-6741 (1989).

PN [2] SEQUENCE FROM N.A.

STRAIN=K12 / MG1655;

RX MEDLINE=92358234; PubMed=1379743;

RA Daniels D.L., Plunkett G. III, Burland V.D., Blattner F.R.;

RT "Analysis of the Escherichia coli genome: DNA sequence of the region

RL from 84.5 to 86.5 minutes.";

RL Science 257:771-778 (1992).

RN [3] SEQUENCE OF 1-17.

RX MEDLINE=93281685; PubMed=8506346;

RA Henzel W.J., Billeci T.M., Stults J.T., Wong S.C., Grimley C.,

RP Watanabe C.;

RT "Identifying proteins from two-dimensional gels by molecular mass

RL searching of peptide fragments in protein sequence databases.";

RL Proc. Natl. Acad. Sci. U.S.A. 90:5011-5015 (1993).

RN [4] PARTIAL SEQUENCE OF 1-11.

STRAIN=K12 / W3110;

RA Frutiger S., Hughes G.J., Pasquali C., Hochstrasser D.F.;

RT Submitted (FEB-1996) to the SWISS-PROT data bank.

RN [5] MUTAGENESIS OF ASP-4.

RX MEDLINE=98326439; PubMed=9661793;

RA Velko V.P., Chebotay D.V., Ovcharova I.V., Gul'ko L.B.;

RT "Protein engineering of uridine phosphorylase from Escherichia coli

RL K-12. I. Cloning and expression of uridine phosphorylase genes from

RN

RT Klebsiella aerogenes and Salmonella typhimurium in E. coli.";  
 RL Bioorg. Khim. 24:381-387 (1998).  
 CC -!- FUNCTION: THE ENZYMES WHICH CATALYZE THE REVERSIBLE PHOSPHORYLOSIS  
 CC OF PYRIMIDINE NUCLEOSIDES ARE INVOLVED IN THE DEGRADATION OF THESE  
 CC COMPOUNDS AND IN THEIR UTILIZATION AS CARBON AND ENERGY SOURCES.  
 CC OR IN THE RESCUE OF PYRIMIDINE BASES FOR NUCLEOTIDE SYNTHESIS.  
 CC -!- CATALYTIC ACTIVITY: Uridine + phosphate = uracil + alpha-D-ribose  
 CC 1-phosphate.  
 CC -!- PATHWAY: Nucleotide and deoxyribonucleotide catabolism.  
 CC -!- SUBUNIT: HOMOHEXAMER. THE HOMOHEXAMER SHOWS 4-FOLD, 6-FOLD OR  
 CC 8-FOLD SYMMETRY.  
 CC -!- SIMILARITY: BELONGS TO THE PNP/UDP FAMILY 1 OF PHOSPHORYLASES.  
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; X15689; CAA33724.1; -

DR EMBL; M87049; AAA67626.1; -

DR EMBL; AE000459; AAC76834.1; -

DR PIR; S05491; S05491.

DR PDB; 1K3F; 10-OCT-01.

DR PDB; 1LX7; 07-JAN-03.

DR EcoGene; EG11045; udp.

DR InterPro; IPR000845; PNP\_UDP.

DR Pfam; PF01048; PNP\_UDP\_1; 1.

DR PROSITE; PS01232; PNP\_UDP\_1; 1.

KW Transferase; Glycosyltransferase; Complete proteome; 3D-structure.

FT INIT MET 0

SQ SEQUENCE 252 AA; 27028 MW; 4D8FDB45A83DD88D CRC64;

Query Match 76.5%; Score 13; DB 1; Length 252;

Best Local Similarity 50.0%; Pred. No. 5.3e+02;

Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7

Db 197 ESATLL 202

RESULT 38

UDP\_SALTY

ID UDP\_SALTY STANDARD; PRT; 252 AA.

AC O33808; O08432; Q9L6M8;

DT 15-JUL-1998 (Rel. 36, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 15-SEP-2003 (Rel. 42, Last annotation update)

DE Uridine phosphorylase (EC 2.4.2.3) (UDPase).

GN UDP OR STM3968 OR STM1.21 OR STY3591 OR T3329.

OS Salmonella typhimurium, and

OS Salmonella typhi.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

OC Enterobacteriaceae; Salmonella.

OX NCBI\_TaxID=602, 601;

RN [1] SEQUENCE FROM N.A.

STRAIN=LT2;

RX SPECIES=S.typhimurium; STRAIN=LT2;

RA Errais L.L., Ukhobotina L.S., Eremina S.Y., Evdokimova A.A.,

RT Mironov A.S.;

RT "Structure and expression of the gene encoding uridine phosphorylase

RL (udp) in Salmonella typhimurium.";

RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.

RN [2] SEQUENCE FROM N.A.

STRAIN=LT2;

RX SPECIES=S.typhimurium; STRAIN=LT2;

RA MEDLINE=98326439; PubMed=9661793;

RA Vaiko V.P., Chebotay D.V., Ovcharova I.V., Gul'ko L.B.;

RT "Protein engineering of uridine phosphorylase from Escherichia coli

RL K-12. I. Cloning and expression of uridine phosphorylase genes from

RN





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2003, 14:14:33 ; Search time 34 Seconds  
(without alignments)  
53.128 Million cell updates/sec

Title: US-09-620-840C-2

Perfect score: 17

Sequence: 1 EXXXLL 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

830525 seqs, 258052604 residues

Minimum number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_23:\*

- 1: sp archaea:\*
- 2: sp bacteria:\*
- 3: sp fungi:\*
- 4: sp human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phase:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_rvirus:\*
- 16: sp\_bacteriap:\*
- 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	ID	Description
1	13	76.5	33	Q9NGN0 strongyloce
2	13	76.5	44	Q90634 baboon hecp
3	13	76.5	52	Q16340 homo sapien
4	13	76.5	53	P88734 human immun
5	13	76.5	54	Q85513 feline sarc
6	13	76.5	63	Q8F1D3 leptospira
7	13	76.5	66	Q1RC2 borna disea
8	13	76.5	67	Q42118 arabidopsis
9	13	76.5	67	Q1RC0 borna disea
10	13	76.5	67	Q1RCB borna disea
11	13	76.5	67	Q1RCB borna disea
12	13	76.5	67	Q1RC1 borna disea
13	13	76.5	67	Q1RCB borna disea
14	13	76.5	67	Q1RCB borna disea
15	13	76.5	85	Q1RCB borna disea
16	13	76.5	91	Q9C9V7 arabidopsis

17	13	76.5	94	3	Q8TFH9	Q8TFH9 schizosacch
18	13	76.5	98	16	Q9RWY3	Q9RWY3 deinococcus
19	13	76.5	102	16	Q93GM9	Q93GM9 salmonella
20	13	76.5	105	16	Q92P93	Q92P93 rhizobium m
21	13	76.5	106	2	Q956M6	Q956M6 chloroflexu
22	13	76.5	107	16	Q9K0C9	Q9K0C9 neisseria m
23	13	76.5	107	16	Q9K0C9	Q9K0C9 neisseria m
24	13	76.5	108	4	Q9NTW6	Q9NTW6 homo sapien
25	13	76.5	110	13	Q8JIU2	Q8JIU2 rana escul
26	13	76.5	110	13	Q8JIU0	Q8JIU0 rana sp. ch
27	13	76.5	114	16	Q9PFMS	Q9PFMS xylella fas
28	13	76.5	115	9	Q94MY5	Q94MY5 haemophilus
29	13	76.5	118	17	Q57918	Q57918 pyrococcus
30	13	76.5	121	16	Q9JY81	Q9JY81 neisseria m
31	13	76.5	121	17	Q9YFQ4	Q9YFQ4 aeropyrum p
32	13	76.5	123	2	Q9EUC1	Q9EUC1 corynebacte
33	13	76.5	124	11	Q8C657	Q8C657 mus musculu
34	13	76.5	126	17	Q8ZXR0	Q8ZXR0 pyrobaculum
35	13	76.5	127	16	Q8FRZ2	Q8FRZ2 corynebacte
36	13	76.5	128	11	Q9CZ88	Q9CZ88 mus musculu
37	13	76.5	128	11	Q9DCQ8	Q9DCQ8 mus musculu
38	13	76.5	129	2	Q30799	Q30799 methyllobact
39	13	76.5	129	17	Q9YD23	Q9YD23 aeropyrum p
40	13	76.5	136	10	Q9LYY0	Q9LYY0 arabidopsis
41	13	76.5	138	16	Q9NTX4	Q9NTX4 corynebacte
42	13	76.5	139	5	Q9BH15	Q9BH15 globodera p
43	13	76.5	140	10	Q8GY72	Q8GY72 arabidopsis
44	13	76.5	140	16	Q9EYD9	Q9EYD9 escherichia
45	13	76.5	140	16	Q8X314	Q8X314 escherichia

## ALIGNMENTS

### RESULT 1

Q9NGN0 PRELIMINARY; PRT; 33 AA.

AC Q9NGN0; 01-OCT-2000 (TREMBLrel. 15, Created)

DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)

DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

DE Myosin IX (Fragment).

OS Strongylocentrotus purpuratus (Purple sea urchin).

OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;

OC Echinozoa; Euechinozoa; Echinacea; Echinoida; Strongylocentrotidae;

OC Strongylocentrotus.

OX NCBI\_TaxID=7668;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=20440049; PubMed=10984411;

RA Sitotkin V., Seipel S., Krendel M., Bondar E.M.;

RT "Characterization of sea urchin unconventional myosins and analysis of

RT their patterns of expression during early embryogenesis.";

RL Mol. Reprod. Dev. 57:111-126(2000).

DR EMBL; AF247511; AAF71723.1; -

DR InterPro; IPR001609; myosin\_head.

DR ProDom; PD000355; myosin\_head; 1.

FT NON TER 1 1

FT NON TER 33 33

SQ SEQUENCE 33 AA; 3406 MW; 85498A96F8914B62 CRC64;

Query Match 76.5%; Score 13; DB 5; Length 33;  
Best Local Similarity 50.0%; Pred. No. 5.2e+02;  
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Oy 2 EXXXLL 7  
Db 1 ESTILL 6

### RESULT 2

Q90634 PRELIMINARY; PRT; 44 AA.

AC O90634;  
 DT 01-NOV-1998 (TrEMBLrel. 08, Created)  
 DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)  
 DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)  
 DE Glycoprotein gp (Fragment).  
 GN US6.  
 OS Baboon herpesvirus 2.  
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
 OC Alphaherpesvirinae; Simplexvirus.  
 OX NCBI\_TaxID=36347;  
 RN [1]  
 RC SEQUENCE FROM N.A.  
 RC STRAIN=8951.  
 RA Eberle R., Smith A., Black D.;  
 RT "Molecular evidence for distinct genotypes of monkey B virus  
 (Herpesvirus simiae) which are related to the Macaque host species.";  
 RL J. Virol. 0:0-0(1998).  
 DR EMBL; AF082810; AAC34107.1; -.  
 FT NON TER 44 44  
 SO SEQUENCE 44 AA; 4427 MW; B3E8DCD1C70BF4EF CRC64;  
 Query Match 76.5%; Score 13; DB 12; Length 44;  
 Best Local Similarity 50.0%; Pred. No. 6.9e+02;  
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 QY 2 EXXXLL 7  
 Db 5 EAAALL 10  
 RESULT 3  
 Q16340  
 ID Q16340 PRELIMINARY; PRT; 52 AA.  
 AC Q16340;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE Phosphodiesterase beta-subunit protein (Fragment).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RC SEQUENCE FROM N.A.  
 RX MEDLINE=95322980; PubMed=7599633;  
 RA Bayes M., Giordano M., Balcells S., Grinberg D., Vilagellu L.,  
 RA Martinez I., Ayuso C., Benitez J., Ramos-Arroyo M.A., Chivelet P.,  
 et al;  
 RT "Homologous tandem duplication within the gene encoding the beta-  
 subunit of rod phosphodiesterase as a cause for autosomal recessive  
 K. retinitis pigmentosa".  
 RL Hum. Mutat. 5:228-234(1995).  
 DR EMBL; S78008; AAD14270.1; -.  
 FT NON TER 1 1  
 SO SEQUENCE 52 AA; 6124 MW; 4FC1A6A8C9DAEF94 CRC64;  
 Query Match 76.5%; Score 13; DB 4; Length 52;  
 Best Local Similarity 50.0%; Pred. No. 8e+02;  
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 QY 2 EXXXLL 7  
 Db 1 ESTALL 6  
 RESULT 4  
 P88734  
 ID P88734 PRELIMINARY; PRT; 53 AA.  
 AC P88734;  
 DT 01-MAY-1997 (TrEMBLrel. 03, Created)  
 DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
 DE Nef protein (Negative factor) (27 kDa protein) (Fragment).  
 GN Nef protein (Negative factor) (27 kDa protein) (Fragment).  
 OS Leptospira interrogans.

GN NEF.  
 OS Human immunodeficiency virus 1.  
 OC Viruses; Retrod viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Quinones-Mateu M.E., Domingo E.;  
 RT "Point mutant frequency and intrasubtype B recombination lead  
 RT Venezuelan HIV-1 genetic diversities".  
 RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING  
 CC ACTIVITIES. IT SEEMS TO DOWN-REGULATE THE CD4 (T4) ANTIGEN (BY  
 CC SIMILARITY).  
 DR EMBL; U81459; AAB39163.1; -.  
 DR InterPro; IPR001558; HIV\_Nef.  
 DR Pfam; PF00469; P-protein; 1.  
 DR ProDom; PD000031; HIV\_Nef; 1.  
 KW AIDS; GTP-binding; Lipoprotein; Myristate.  
 FT NON TER 1 1  
 SO SEQUENCE 53 AA; 6295 MW; CBB1FED2B7C3CB05 CRC64;  
 Query Match 76.5%; Score 13; DB 15; Length 53;  
 Best Local Similarity 50.0%; Pred. No. 8.1e+02;  
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 QY 2 EXXXLL 7  
 Db 7 ETSSLL 12  
 RESULT 5  
 Q85513  
 ID Q85513 PRELIMINARY; PRT; 54 AA.  
 AC Q85513;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
 DE (SM) v-fms 3' recombination site RNA (Fragment).  
 OS Feline sarcoma virus.  
 OC Viruses; Retrod viruses; Retroviridae; Mammalian type C retroviruses.  
 OX NCBI\_TaxID=11772;  
 RN [1]  
 RC SEQUENCE FROM N.A.  
 RX MEDLINE=86308237; PubMed=3018286;  
 RA Besmer P., Lader E., George P.C., Bergold P.J., Qiu P.-H.,  
 RA Zuckerman E.E., Hardy W.D.;  
 RT "A new acute transforming feline retrovirus with fms homology  
 RT specifies a C-terminally truncated version of the c-fms protein that  
 RT is different from SM-Feline sarcoma virus v-fms protein.";  
 RL J. Virol. 60:194-203(1986).  
 DR EMBL; M14291; AAA43044.1; -.  
 FT NON TER 1 1  
 SO SEQUENCE 54 AA; 5583 MW; 7C116DDDCARCFB93 CRC64;  
 Query Match 76.5%; Score 13; DB 15; Length 54;  
 Best Local Similarity 50.0%; Pred. No. 8.3e+02;  
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 QY 2 EXXXLL 7  
 Db 49 ESSSLL 54  
 RESULT 6  
 Q8FID3  
 ID Q8FID3 PRELIMINARY; PRT; 63 AA.  
 AC Q8FID3;  
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Hypothetical protein.  
 GN LA3203.  
 OS Leptospira interrogans.

OC Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospira.  
 RN NCBI\_TaxID=173;  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=56601 / Serogroup Icterohaemorrhagiae / Serovar lai,  
 RA Ren S.;  
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AB011482; AAN50401.1; -;  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 63 AA; 7252 MW; 616049C944C3DF26 CRC64;

Query Match 76.5%; Score 13; DB 16; Length 63;  
 Best Local Similarity 50.0%; Pred. No. 9.6e+02;  
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7  
 Db 24 ETTLL 29

RESULT 8  
 Q42118 PRELIMINARY; PRT; 67 AA.  
 ID Q42118; (TREMELrel. 01, Created)  
 DT 01-NOV-1996 (TREMELrel. 09, Last sequence update)  
 DT 01-JAN-1999 (TREMELrel. 23, Last annotation update)  
 DE Prolin-rich protein APG (Fragment).  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia; TISSUE=Seedling;  
 RA Hofte H.;  
 RL Submitted (SEP-1993) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; Z26539; CAA81310.1; -;

OC Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospira.  
 RN NCBI\_TaxID=173;  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=56601 / Serogroup Icterohaemorrhagiae / Serovar lai,  
 RA Ren S.;  
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AB011482; AAN50401.1; -;  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 63 AA; 7252 MW; 616049C944C3DF26 CRC64;

Query Match 76.5%; Score 13; DB 16; Length 63;  
 Best Local Similarity 50.0%; Pred. No. 9.6e+02;  
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7  
 Db 24 ETTLL 29

RESULT 8  
 Q42118 PRELIMINARY; PRT; 67 AA.  
 ID Q42118; (TREMELrel. 01, Created)  
 DT 01-NOV-1996 (TREMELrel. 09, Last sequence update)  
 DT 01-JAN-1999 (TREMELrel. 23, Last annotation update)  
 DE Prolin-rich protein APG (Fragment).  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia; TISSUE=Seedling;  
 RA Hofte H.;  
 RL Submitted (SEP-1993) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; Z26539; CAA81310.1; -;

DR InterPro; IPR001087; Lipase GDSL.  
 DR Pfam; PF00657; Lipase GDSL; 1.  
 FT NON\_TER 1  
 SQ SEQUENCE 67 AA; 6902 MW; FDCAD1ES3E3E6A98 CRC64;

Query Match 76.5%; Score 13; DB 10; Length 67;  
 Best Local Similarity 50.0%; Pred. No. 1e+03;  
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7  
 Db 18 ETTSL 23

RESULT 9  
 Q91RC0 PRELIMINARY; PRT; 67 AA.  
 ID Q91RC0  
 AC Q91RC0; (TREMELrel. 19, Created)  
 DT 01-DEC-2001 (TREMELrel. 19, Last sequence update)  
 DT 01-DEC-2001 (TREMELrel. 19, Last annotation update)  
 DE P40 (Fragment).  
 OS Borna disease virus (BDV).  
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales; Bornaviridae;  
 OC Bornavirus.  
 OX NCBI\_TaxID=12455;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=25;  
 RA Zientara S.; Dauphin G.; Legay V.; Hammoui S.; Smondack S.;  
 RA Saillieu C.;  
 RT "Evidence of Borna disease virus genome detection in French domestic  
 animals and in foxes (Vulpes vulpes).";  
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF374602; AAK54625.1; -;  
 FT NON\_TER 1  
 FT NON\_TER 67  
 SQ SEQUENCE 67 AA; 7337 MW; 347D89F06C14C33B CRC64;

Query Match 76.5%; Score 13; DB 12; Length 67;  
 Best Local Similarity 50.0%; Pred. No. 1e+03;  
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7  
 Db 57 ETATLL 62

RESULT 10  
 Q91RB7 PRELIMINARY; PRT; 67 AA.  
 ID Q91RB7  
 AC Q91RB7; (TREMELrel. 19, Created)  
 DT 01-DEC-2001 (TREMELrel. 19, Last sequence update)  
 DT 01-DEC-2001 (TREMELrel. 19, Last annotation update)  
 DE P40 (Fragment).  
 OS Borna disease virus (BDV).  
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales; Bornaviridae;  
 OC Bornavirus.  
 OX NCBI\_TaxID=12455;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=204;  
 RA Zientara S.; Dauphin G.; Legay V.; Hammoui S.; Smondack S.;  
 RA Saillieu C.;  
 RT "Evidence of Borna disease virus genome detection in French domestic  
 animals and in foxes (Vulpes vulpes).";  
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF374605; AAK54628.1; -;  
 FT NON\_TER 1  
 FT NON\_TER 67  
 SQ SEQUENCE 67 AA; 7337 MW; 347D89F06C14C33B CRC64;

Query Match 76.5%; Score 13; DB 12; Length 67;  
 Best Local Similarity 50.0%; Pred. No. 1e+03;  
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7  
 Db 57 ETATLL 62

## RESULT 11

ID Q91RB6 PRELIMINARY; PRT; 67 AA.  
 AC Q91RB6; (1)  
 DT 01-DEC-2001 (TReMBLrel. 19, Created)  
 DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)  
 DE P40 (Fragment)  
 OS Borna disease virus (BDV)  
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales; Bornaviridae;  
 OC Bornavirus  
 NCBI\_TaxID=12455;  
 (1)  
 kF SEQUENCE FROM N.A.  
 RC STRAIN=175;  
 RA Zientara S., Dauphin G., Legay V., Hammoui S., Smondack S.,  
 RA Sailleau C.;  
 RT "Evidence of Borna disease virus genome detection in French domestic  
 animals and in foxes (Vulpes vulpes).";  
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF374606; AAK54629.1;  
 FT NON\_TER 1  
 FT NON\_TER 67  
 SQ SEQUENCE 67 AA; 7337 MW; 347D89F06C14C33B CRC64;

Query Match 76.5%; Score 13; DB 12; Length 67;  
 Best Local Similarity 50.0%; Pred. No. 1e+03;  
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7  
 Db 57 ETATLL 62

## RESULT 12

ID Q91RC1 PRELIMINARY; PRT; 67 AA.  
 AC Q91RC1; (1)  
 DT 01-DEC-2001 (TReMBLrel. 19, Created)  
 DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)  
 DE P40 (Fragment)  
 OS Borna disease virus (BDV)  
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales; Bornaviridae;  
 OC Bornavirus  
 NCBI\_TaxID=12455;  
 (1)  
 kF SEQUENCE FROM N.A.  
 RC STRAIN=18;  
 RA Zientara S., Dauphin G., Legay V., Hammoui S., Smondack S.,  
 RA Sailleau C.;  
 RT "Evidence of Borna disease virus genome detection in French domestic  
 animals and in foxes (Vulpes vulpes).";  
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF374601; AAK54624.1;  
 FT NON\_TER 1  
 FT NON\_TER 67  
 SQ SEQUENCE 67 AA; 7337 MW; 347D89F06C14C33B CRC64;

Query Match 76.5%; Score 13; DB 12; Length 67;  
 Best Local Similarity 50.0%; Pred. No. 1e+03;  
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7

Db 57 ETATLL 62

## RESULT 13

ID Q91RB9 PRELIMINARY; PRT; 67 AA.  
 AC Q91RB9; (1)  
 DT 01-DEC-2001 (TReMBLrel. 19, Created)  
 DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)  
 DE P40 (Fragment)  
 OS Borna disease virus (BDV)  
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales; Bornaviridae;  
 OC Bornavirus  
 NCBI\_TaxID=12455;  
 (1)  
 kF SEQUENCE FROM N.A.  
 RC STRAIN=63;  
 RA Zientara S., Dauphin G., Legay V., Hammoui S., Smondack S.,  
 RA Sailleau C.;  
 RT "Evidence of Borna disease virus genome detection in French domestic  
 animals and in foxes (Vulpes vulpes).";  
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF374603; AAK54626.1;  
 FT NON\_TER 1  
 FT NON\_TER 67  
 SQ SEQUENCE 67 AA; 7337 MW; 347D89F06C14C33B CRC64;

Query Match 76.5%; Score 13; DB 12; Length 67;  
 Best Local Similarity 50.0%; Pred. No. 1e+03;  
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7  
 Db 57 ETATLL 62

## RESULT 14

ID Q91RB8 PRELIMINARY; PRT; 67 AA.  
 AC Q91RB8; (1)  
 DT 01-DEC-2001 (TReMBLrel. 19, Created)  
 DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)  
 DE P40 (Fragment)  
 OS Borna disease virus (BDV)  
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales; Bornaviridae;  
 OC Bornavirus  
 NCBI\_TaxID=12455;  
 (1)  
 kF SEQUENCE FROM N.A.  
 RC STRAIN=198;  
 RA Zientara S., Dauphin G., Legay V., Hammoui S., Smondack S.,  
 RA Sailleau C.;  
 RT "Evidence of Borna disease virus genome detection in French domestic  
 animals and in foxes (Vulpes vulpes).";  
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF374604; AAK54627.1;  
 FT NON\_TER 1  
 FT NON\_TER 67  
 SQ SEQUENCE 67 AA; 7337 MW; 347D89F06C14C33B CRC64;

Query Match 76.5%; Score 13; DB 12; Length 67;  
 Best Local Similarity 50.0%; Pred. No. 1e+03;  
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7  
 Db 57 ETATLL 62

## RESULT 15



Q9YNV8  
ID Q9YNV8 PRELIMINARY; PRT; 85 AA.  
AC Q9YNV8  
DT 01-JUN-1999 (TREMELrel. 10, Created)  
DT 01-MAY-1999 (TREMELrel. 10, Last sequence update)  
DT 01-JUN-2001 (TREMELrel. 17, Last annotation update)  
DE AL4.  
GN AL4.  
OS Sida golden mosaic virus.  
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
OC NCBI\_TaxID=51034;  
RN [1] SEQUENCE FROM N.A.  
RP STRAIN=DNA-A1;  
RC Yung Z., Davis M.J.;  
RA "Complete nucleotide sequences of two sida golden mosaic viruses in  
RT South Florida."  
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.  
EMBL: U77964; AAD10474.1;  
InterPro; IPR002488; Gemini\_C4.  
Pfam; PF01492; Gemini\_C4; 1.  
SQ SEQUENCE 85 AA; 9677 MW; D45D988F3ADC05F7 CRC64;

Query Match 76.5%; Score 13; DB 12; Length 85;  
Best Local Similarity 50.0%; Pred. No. 1.3e+03;  
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7  
DB 73 EASLL 78

## RESULT 16

ID Q9C9V7 PRELIMINARY; PRT; 91 AA.  
AC Q9C9V7  
DT 01-JUN-2001 (TREMELrel. 17, Created)  
DT 01-JUN-2001 (TREMELrel. 17, Last sequence update)  
DT 01-JUN-2001 (TREMELrel. 17, Last annotation update)  
DE Hypothetical 9.9 kDa protein.  
GN T23K23.24.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.  
OC NCBI\_TaxID=3702;  
P- [1]  
SEQUENCE FROM N.A.  
STRAIN=sv. Columbia;  
MEDLINE=21016719; PubMed=11130712;  
Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,  
White O., Alonso J., Altafi H., Araujo R., Bowman C.L., Brooks S.Y.,  
Buehler E., Chan O., Chen H., Cheuk R.P., Chin C.W.,  
Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,  
Dunn P., Egtu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y.,  
Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,  
Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,  
Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,  
Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marziani A.,  
Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,  
Militischer J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,  
Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,  
Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,  
Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,  
Utterback T., Van Aken S., Vaysberg M., Vayatskaia V.S., Walker M.,  
Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;  
"Sequence and analysis of chromosome 1 of the plant Arabidopsis  
thaliana";  
Nature 408:816-820 (2000).  
EMBL; AC012563; AAG52013.1;  
Hypothetical protein.  
SQ SEQUENCE 91 AA; 9940 MW; 80C750FA20EBF8D2 CRC64;

Query Match 76.5%; Score 13; DB 10; Length 91;  
Best Local Similarity 50.0%; Pred. No. 1.3e+03;  
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7  
DB 72 EASLL 77

## RESULT 17

ID Q8TFH9 PRELIMINARY; PRT; 94 AA.  
AC Q8TFH9  
DT 01-JUN-2002 (TREMELrel. 21, Created)  
DT 01-JUN-2002 (TREMELrel. 21, Last sequence update)  
DT 01-JUN-2002 (TREMELrel. 21, Last annotation update)  
DE Putative cytochrome c oxidase assembly protein, putative requirement  
DE for heme A synthesis (Fragment).  
GN COX11.  
OS Schizosaccharomyces pombe (Fission yeast).  
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
OC Schizosaccharomycetes.  
OX NCBI\_TaxID=4896;  
RN [1] SEQUENCE FROM N.A.  
RP STRAIN=972h-;  
RC O'Neil S., Harris D., Wood V., Rajandream M.A., Barrell B.G.;  
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AL691405; CAD27494.1;  
FT NON TER 94  
SQ SEQUENCE 94 AA; 11137 MW; 6426630F09053CB7 CRC64;

Query Match 76.5%; Score 13; DB 3; Length 94;  
Best Local Similarity 50.0%; Pred. No. 1.4e+03;  
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7  
DB 65 EASLL 70

## RESULT 18

ID Q9RWY3 PRELIMINARY; PRT; 98 AA.  
AC Q9RWY3  
DT 01-MAY-2000 (TREMELrel. 13, Created)  
DT 01-MAY-2000 (TREMELrel. 13, Last sequence update)  
DT 01-MAR-2002 (TREMELrel. 20, Last annotation update)  
DE Hypothetical protein DR0532.  
GN DR0532.  
OS Deinococcus radiodurans.  
OC Bacteria; Deinococcus-Thermus; Deinococci; Deinococcales;  
OC Deinococcaceae; Deinococcus.  
OX NCBI\_TaxID=1299;  
RN [1] SEQUENCE FROM N.A.  
RP STRAIN=R1;  
RC MEDLINE=20036896; PubMed=10567266;  
RA White O., Eisen J.A., Heidelberg J.P., Hickey E.K., Peterson J.D.,  
Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,  
Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,  
Vamathevan J.J., Lam P., McDonald L., Utterback T., Zaleski C.,  
Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,  
Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,  
Fraser C.M.;  
"Genome sequence of the radioresistant bacterium Deinococcus  
radiodurans R1";  
Science 286:1571-1577 (1999).  
EMBL; AE001911; AAF10113.1;  
DR TIGR; DR0532;  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 98 AA; 10926 MW; 135DFCA22BB2A639 CRC64;

Query Match 76.5%; Score 13; DB 16; Length 98;  
 Best Local Similarity 50.0%; Pred. No. 1.4e+03;  
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7  
 DB 43 EYALL 49

## RESULT 19

Q93GM9 PRELIMINARY; PRT; 102 AA.  
 AC Q93GM9; (TREMBlrel. 19, Created)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
 DE Conjugative transfer.  
 GN PSL7093  
 OS Salmonella typhimurium.  
 Plasmid pSLT.  
 Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 Enterobacteriaceae; Salmonella.  
 NCBI\_TaxID=602;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=LT2 / SGSC1412 / ATCC 700720;  
 RX MEDLINE=21534948; PubMed=11677609;  
 RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,  
 RA Courtney L., Porwollik S., Ali J., Dante M., Du P., Hou S., Layman D.,  
 RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,  
 RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,  
 RA Waterston R., Wilson R.K.;  
 RT "Complete genome sequence of Salmonella enterica serovar Typhimurium  
 LT2.";  
 RL Nature 413:852-856(2001).  
 DR EMBL; AE006471; AAL23496.1; --  
 KW Plasmid; Complete proteome.  
 SQ SEQUENCE 102 AA; 11958 MW; CA113B13A6A19344 CRC64;

Query Match 76.5%; Score 13; DB 16; Length 102;  
 Best Local Similarity 50.0%; Pred. No. 1.5e+03;  
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7  
 DB 86 ESSLL 91

## JLT 20

Q92P93 PRELIMINARY; PRT; 105 AA.  
 AC Q92P93; (TREMBlrel. 19, Created)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
 DE Putative transcription regulator protein.  
 GN R01890 OR SMC04220.  
 OS Rhizobium meliloti (Sinorhizobium meliloti).  
 Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
 Rhizobiaceae; Sinorhizobium.  
 NCBI\_TaxID=382;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=1021;  
 RX MEDLINE=21396507; PubMed=11481430;  
 RA Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J.,  
 RA Bolstad P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,  
 RA Godrie T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masuy D.,  
 RA Pohl T., Porretelle D., Puehler A., Purnelle B., Ransperger U.,  
 RA Renard C., Thebault P., Vandenberg M., Weidner S., Galibert F.;  
 RT "Analysis of the chromosome sequence of the legume symbiont  
 Sinorhizobium meliloti strain 1021.";

Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).  
 DR EMBL; AL591788; CAC46469.1; --  
 DR InterPro; IPR001845; HTH\_ArsR.  
 DR Pfam; PF01022; HTH\_5; 1.  
 DR PRINTS; PRO0778; HTHARSR.  
 DR SMART; SMO0418; HTH\_ArsR; 1.  
 KW Complete proteome.  
 SQ SEQUENCE 105 AA; 11880 MW; FE9F38BB18F6B22F CRC64;

Query Match 76.5%; Score 13; DB 16; Length 105;  
 Best Local Similarity 50.0%; Pred. No. 1.5e+03;  
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7  
 DB 20 EYALL 25

## RESULT 21

Q9F6W6 PRELIMINARY; PRT; 106 AA.  
 AC Q9F6W6; (TREMBlrel. 16, Created)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
 DE Hypothetical 11.7 kDa protein.  
 OS Chloroflexus aurantiacus.  
 Bacteria; Chloroflexi; Chloroflexales; Chloroflexaceae; Chloroflexus.  
 NCBI\_TaxID=1108;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20433268; PubMed=10976061;  
 RA Xiong J., Fischer W.M., Inoue K., Nakahara M., Bauer C.E.;  
 RT "Molecular evidence for the early evolution of photosynthesis.";  
 RL Science 289:1724-1730(2000).  
 DR EMBL; AF288461; AAG15219.1; --  
 KW Hypothetical protein.  
 SQ SEQUENCE 106 AA; 11666 MW; B9D117929519ED70 CRC64;

Query Match 76.5%; Score 13; DB 2; Length 106;  
 Best Local Similarity 50.0%; Pred. No. 1.5e+03;  
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7  
 DB 14 ESSLL 19

## RESULT 22

Q9K0C9 PRELIMINARY; PRT; 107 AA.  
 AC Q9K0C9; (TREMBlrel. 15, Created)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
 DE Hypothetical protein NMB0685.  
 GN NMB0685.  
 OS Neisseria meningitidis (serogroup B).  
 Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;  
 OC Neisseriaceae; Neisseria.  
 NCBI\_TaxID=491;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MC58 / Serogroup B;  
 RX MEDLINE=20175755; PubMed=10710307;  
 RA Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,  
 RA Eisen J.A., Kachum K.A., Hood D.W., Peden J.F., Dodson R.J.,  
 RA Nelson W.C., Gwinn M.L., Deboy R., Peterson J.D., Hickey E.K.,  
 RA Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,  
 RA Mason T., Ciecko A., Parksey D.S., Blair E., Ciccione H., Clark E.B.,  
 RA Cotton M.D., Utterback T.R., Khouri H., Qin H., Vamathevan J.,  
 RA Gill J., Scarlato V., Maignani V., Pizzi M., Grandi G., Sun L.,  
 RA Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R., Venter J.C.;

```

RT "Complete genome sequence of Neisseria meningitidis serogroup B strain
RT MCS8.";
RL Science 287:1809-1815(2000).
DR EMBL; AB002422; AAF41103.1; -.
DR TIGR; NMB0685; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 107 AA; 11797 MW; 54F6D701E3610A9F CRC64;

Query Match 76.5%; Score 13; DB 16; Length 107;
Best Local Similarity 50.0%; Pred. No. 1.6e+03;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7
DB 76 ETRALL 81

RESULT 23
CQTVDA PRELIMINARY; PRT; 107 AA.
OS Q9JVD4
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Hypothetical protein NMA0887.
GN NMA0887.
OS Neisseria meningitidis (serogroup A).
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxID=65699;
RN [1]
RC SEQUENCE FROM N.A.
RP STRAIN=22491 / Serogroup A / Serotype 4A;
RX MEDLINE=20222556; PubMed=10761919;
RA Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.,
RA Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
RA Davies R.M., Davis P., Devlin K., Feltham T., Hamlin N., Holroyd S.,
RA Jagels K., Leathers S., Moule S., Mungall K., Quail M.A.,
RA Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,
RA Whitehead S., Spratt B.G., Barrall B.G.;
RT "Complete DNA sequence of a serogroup A strain of Neisseria
RT meningitidis Z2491.";
RL Nature 404:502-506(2000).
DR EMBL; AL162754; CAB84167.1; -.
DR Hypothetical protein; Complete proteome.
SQ SEQUENCE 107 AA; 11799 MW; 54B158FC6CB82C7F CRC64;

Query Match 76.5%; Score 13; DB 16; Length 107;
Best Local Similarity 50.0%; Pred. No. 1.6e+03;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7
DB 76 ETRALL 81

RESULT 24
Q9NTW6 PRELIMINARY; PRT; 108 AA.
ID Q9NTW6
AC Q9NTW6;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE DJ777D9.1.1 (Hypothalamus protein HT013, isoform 1) (Fragment).
GN DJ198K11.1
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RC SEQUENCE FROM N.A.
RP Blakey S.;
RA Blakey S.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.

```

```

DR EMBL; AL110120; CAB75609.1; -.
FT NON_TER 1
FT NON_TER 108 108
SQ SEQUENCE 108 AA; 11771 MW; 5B5435ECC0DA42BB CRC64;

Query Match 76.5%; Score 13; DB 4; Length 108;
Best Local Similarity 50.0%; Pred. No. 1.6e+03;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7
DB 31 ETRALL 36

RESULT 25
Q8JIU2 PRELIMINARY; PRT; 110 AA.
ID Q8JIU2
AC Q8JIU2;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Parvalbumin alpha.
OS Rana esculenta (Edible frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranoidae; Ranidae; Rana.
OX NCBI_TaxID=8401;
RN [1]
RC SEQUENCE FROM N.A.
RP TISSUE=Muscle;
RA Hilger C., Grigioni F., Thill L., Mertens L., Hentges F.;
RT "Severe anaphylaxis upon consumption of fried frog legs. Definition of
RT alpha-parvalbumin as the allergen in cause.";
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ315959; CAC83046.1; -.
DR InterPro; IPR002048; EF-hand.
DR Pfam; PF00036; ehand; 2.
DR ProDom; PD000012; EF-hand; 1.
DR SMART; SM00034; EFh; 2.
DR PROSITE; PS00018; EF HAND; 2.
SQ SEQUENCE 110 AA; 11929 MW; 4083650C4FADF37 CRC64;

Query Match 76.5%; Score 13; DB 13; Length 110;
Best Local Similarity 50.0%; Pred. No. 1.6e+03;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7
DB 82 ETRALL 87

RESULT 26
Q8JIU0 PRELIMINARY; PRT; 110 AA.
ID Q8JIU0
AC Q8JIU0;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Parvalbumin alpha.
OS Rana sp. CH-2001.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranoidae; Ranidae; Rana.
OX NCBI_TaxID=169395;
RN [1]
RC SEQUENCE FROM N.A.
RP TISSUE=Muscle;
RA Hilger C., Grigioni F., Thill L., Mertens L., Hentges F.;
RT "Severe anaphylaxis upon consumption of fried frog legs. Definition of
RT alpha-parvalbumin as the allergen in cause.";
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ315960; CAC83047.1; -.
DR InterPro; IPR002048; EF-hand.
DR Pfam; PF00036; ehand; 2.
DR ProDom; PD000012; EF-hand; 1.

```

```

DR SMART; SMO0054; EPH; 2.
DR PROSITE; PS00018; EF HAND; 2.
SQ SEQUENCE 110 AA; 12002 MW; FF6193B5237EAB04 CRC64;

Query Match 76.5%; Score 13; DB 13; Length 110;
Best Local Similarity 50.0%; Pred. No. 1.7e+03;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7
DB 82 ETALL 87

RESULT 27
Q9PFMS PRELIMINARY; PRT; 114 AA.
AC Q9PFMS;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
DR EMBL; AY0632.
SQ SEQUENCE 114 AA; 12348 MW; 7257B73F96D57FF3 CRC64;

Query Match 76.5%; Score 13; DB 16; Length 114;
Best Local Similarity 50.0%; Pred. No. 1.7e+03;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7
DB 49 EAAALL 54

RESULT 28
Q94MY5 PRELIMINARY; PRT; 115 AA.
AC Q94MY5;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
DR EMBL; AY027935; AAK37809.1;
SQ SEQUENCE 115 AA; 13075 MW; 390F5EF7227741D2 CRC64;

Query Match 76.5%; Score 13; DB 9; Length 115;
Best Local Similarity 50.0%; Pred. No. 1.7e+03;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7
DB 35 ETALL 40

RESULT 29
Q57918 PRELIMINARY; PRT; 118 AA.
AC Q57918;
DT 01-AUG-1998 (TReMBLrel. 07, Created)
DT 01-AUG-1998 (TReMBLrel. 07, Last sequence update)
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
DR EMBL; PH0179.
SQ SEQUENCE 118 AA; 12438 MW; F29C15211DC38997 CRC64;

Query Match 76.5%; Score 13; DB 17; Length 118;
Best Local Similarity 50.0%; Pred. No. 1.7e+03;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7
DB 46 ETALL 51

RESULT 30
Q9JY81 PRELIMINARY; PRT; 121 AA.
AC Q9JY81;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)

```

```

DE Hypothetical protein NMB1699.
GN NMB1699.
OS Neisseria meningitidis (serogroup B).
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OK NCBI_TaxID=491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MC58 / Serogroup B;
RX MEDLINE=20175755; PubMed=10710307;
RA Tetelin H., Saunders N.J., Heideberg J., Jeffries A.C., Nelson K.E.,
RA Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,
RA Nelson W.C., Gwinn M.L., DeBoy R., Peterson J.D., Hickey E.K.,
RA Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
RA Mason T., Ciecko A., Parkey D.S., Blair E., Ciftone H., Clark E.B.,
RA Cotton M.D., Uterback T.R., Khouri H., Qin H., Vamathevan J.,
RA Gill J., Scariato V., Massignani V., Pizzo M., Grandi G., Sun L.,
RA Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R., Venter J.C.;
P "Complete genome sequence of Neisseria meningitidis serogroup B strain
MC58."
Science 287:1809-1815(2000).
DR EMBL; AE002520; AAF42047.1; -.
DR TIGR; NMB1699; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 121 AA; 13195 MW; BBDCA07A3BF533EF CRC64;

Query Match 76.5%; Score 13; DB 16; Length 121;
Best Local Similarity 50.0%; Pred. No. 1.8e+03;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7
DB 20 EAASLL 25

RESULT 31
QYFQ4 PRELIMINARY; PRT; 121 AA.
AC QYFQ4;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DE 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical protein APE0195.
GN APE0195.
OS Aeropyrum pernix.
OC Archaea; Crenarchaeota; Thermoprotei; Desulfurococcaceae;
OC Desulfurococcaceae; Aeropyrum.
OK NCBI_TaxID=56636;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K1;
RX MEDLINE=99310339; PubMed=10382966;
RA Kawarabayashi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,
RA Jinno K., Takahashi M., Sekine M., Baba S.-I., Ankaï A., Kosugi H.,
RA Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,
RA Yamazaki J., Kushiida N., Oguchi A., Aoki K.-I., Kubota K.,
RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.;
RT "Complete genome sequence of an aerobic hyper-thermophilic
RT crenarchaeon, Aeropyrum pernix K1."
RL DNA Res. 6:83-101(1999).
DR EMBL; AP000058; BAA79107.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 121 AA; 13690 MW; 1AD93093A8837E4 CRC64;

Query Match 76.5%; Score 13; DB 17; Length 121;
Best Local Similarity 50.0%; Pred. No. 1.8e+03;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7
DB 34 ETASLL 39

```

```

RESULT 32
Q9EUC1 PRELIMINARY; PRT; 123 AA.
AC Q9EUC1;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE Hypothetical 13.5 kDa protein.
GN ORF55.
OS Corynebacterium equi (Rhodococcus equi).
OC Plasmid PRAT701 (p33701), and Plasmid virulence.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Nocardiaceae; Rhodococcus.
OK NCBI_TaxID=43767;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC33701; PLASMID=PRAT701 (p33701);
RA Takai S., Sekizaki T., Kakuda T., Nakamura M., Suzuki K., Ogino N.;
RL Submitted (FEB-2000) to the EMBL/GenBank/DBSJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC33701, and 103; PLASMID=PRAT701 (p33701), and virulence;
RX PubMed=11083803;
RA Takai S., Hines S.A., Sekizaki T., Nicholson V.M., Alperin D.A.,
RA Osaki M., Takamatsu D., Nakamura M., Suzuki K., Ogino N., Kakuda T.,
RA Dan H., Prescott J.F.;
RT "DNA sequence and comparison of virulence plasmids from Rhodococcus
RT equi ATCC 33701 and 103."
RL Infect. Immun. 68:6840-6847(2000).
DR EMBL; AF001204; BAB16665.1; -.
DR EMBL; AF116907; AAG21759.1; -.
DR InterPro; IPR001387; HTH_3.
DR Pfam; PF01381; HTH_3; 1.
DR SMART; SM00530; HTH_XRE; 1.
KW Hypothetical protein; Plasmid.
SQ SEQUENCE 123 AA; 13516 MW; 0E840DD82B962DA CRC64;

Query Match 76.5%; Score 13; DB 2; Length 123;
Best Local Similarity 50.0%; Pred. No. 1.8e+03;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7
DB 25 ETASLL 30

RESULT 33
Q8C657 PRELIMINARY; PRT; 124 AA.
AC Q8C657;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OK NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Head;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium.
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573(2002).
DR EMBL; AK076534; BAC36383.1; -.
KW Hypothetical protein.
SQ SEQUENCE 124 AA; 14425 MW; 935656E0C8599E46 CRC64;

Query Match 76.5%; Score 13; DB 11; Length 124;

```

Best Local Similarity 50.0%; Pred. No. 1.8e+03;  
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7  
DB 34 EASTLL 39

## RESULT 34

Q82XRO PRELIMINARY; PRT; 126 AA.  
AC Q82XRO;  
DT 01-MAR-2002 (TRENBLrel. 20, Created)  
DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)  
DE Hypothetical protein PAE1151.  
GN PAE1151.  
OS Pyrobaculum aerophilum.  
OC Archaea; Crenarchaeota; Thermoprotei; Thermoproteales;  
OC Thermoproteaceae; Pyrobaculum.  
[1]  
RC STRAIN=IM2 / ATCC 51768 / DSM 7523;  
RX PubMed:11792869;  
RA Fitz-Gibbon S.T., Ladner H., Kim U.-J., Stetter K.O., Simon M.I.,  
RA Miller J.H.;  
RT "Genome sequence of the hyperthermophilic crenarchaeon Pyrobaculum  
RT aerophilum";  
RL EMBL; AEO09806; AAL63286.1; --  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 126 AA; 14559 MW; E3A2BC07BE2123C9 CRC64;

Query Match 76.5%; Score 13; DB 17; Length 126;  
Best Local Similarity 50.0%; Pred. No. 1.8e+03;  
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7  
DB 111 EASTLL 116

## RESULT 35

Q8FRZ2 PRELIMINARY; PRT; 127 AA.  
AC Q8FRZ2;  
DT 01-MAR-2003 (TRENBLrel. 23, Created)  
DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)  
DE Hypothetical protein.  
GN CE0616.  
OS Corynebacterium efficiens.  
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
OC Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.  
OX NCBI\_TaxID=152794;  
[1]  
RC STRAIN=YS-314 / AJ 12310 / DSM 44549 / JCM 11189;  
RA Kawabayashi Y., Yamazaki J., Hino Y., Kikuchi H., Nakamura Y.,  
RA Ikeo K., Suzuki M., Mashima J., Itoh T., Yamagishi A., Nishio Y.,  
RA Usuda Y., Sugimoto S.;  
RT "The entire genomic sequence of Corynebacterium efficiens YS-314";  
RL Submitted (May-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AP005216; BAC17426.1; --  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 127 AA; 13842 MW; 4D1FEB5E3C1D55B1 CRC64;

Query Match 76.5%; Score 13; DB 16; Length 127;  
Best Local Similarity 50.0%; Pred. No. 1.8e+03;  
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7

Db 13 EASTLL 18

## RESULT 36

Q9CZ88 PRELIMINARY; PRT; 128 AA.  
AC Q9CZ88;  
DT 01-JUN-2001 (TRENBLrel. 17, Created)  
DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)  
DT 01-JUN-2001 (TRENBLrel. 17, Last annotation update)  
DE 0610012A05Rik protein.  
GN 0610012A05Rik.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
[1]  
RC SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Embryo;  
RX MEDLINE=21085660; PubMed=11217851;  
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,  
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,  
RA Schriml L.M., Staubli P., Suzuki R., Tomita M., Wagner L., Washio T.,  
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.P.,  
RA Brownstein M.J., Bult C., Fietcher C., Fujita M., Gariboldi M.,  
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombarts P.,  
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,  
RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,  
RA Hayashizaki Y.;  
RT "Functional annotation of a full-length mouse cDNA collection";  
RL Nature 409:685-690 (2001).  
DR EMBL; AK012871; EAB28525.1; --  
DR MGD; MGI:1914684; 0610012A05Rik.  
SQ SEQUENCE 128 AA; 13429 MW; 829AA670C2F2DD4C CRC64;

Query Match 76.5%; Score 13; DB 11; Length 128;  
Best Local Similarity 50.0%; Pred. No. 1.8e+03;  
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7  
DB 81 EASTLL 86

## RESULT 37

Q9DCQ8 PRELIMINARY; PRT; 128 AA.  
AC Q9DCQ8;  
DT 01-JUN-2001 (TRENBLrel. 17, Created)  
DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)  
DT 01-JUN-2001 (TRENBLrel. 17, Last annotation update)  
DE 0610012A05Rik protein.  
GN 0610012A05Rik.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
[1]  
RC SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Kidney;  
RX MEDLINE=21085660; PubMed=11217851;  
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,

RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,  
 RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
 RA Kuehl P., Lewis S., Mateu Y., Nikiado I., Pasole G., Quackenbush J.,  
 RA Schriml L.M., Straubli P., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,  
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,  
 RA Hayashizaki Y.,  
 RT "Functional annotation of a full-length mouse cDNA collection."  
 RL Nature 409:685-690(2001).  
 RA EMBL, AK02572; BAB22197.1; --  
 MGDI:1914684; 0610012A05Rik.  
 SEQUENCE 128 AA; 13415 MW; 829AA2E6199BFD4C CRC64;

Query Match 76.5%; Score 13; DB 11; Length 128;  
 Best Local Similarity 50.0%; Pred. No. 1.8e+03;  
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7  
 |  
 |  
 |  
 DB 81 EAATLL 86

RESULT 38  
 Q30799  
 ID Q30799 PRELIMINARY; PRT; 129 AA.  
 AC Q30799;  
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)  
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE Methanol oxidation genes, mxaB, mxaH, mxaB, and pmi-like genes.  
 OS Methylobacterium extorquens.  
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
 OC Methylobacteriaceae; Methylobacterium.  
 OX NCBI\_TaxID=408;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=AM1;  
 PV MEDLINE=98156126; PubMed=9495022;  
 RA Springer A.L., Auman A.J., Lidstrom M.E.;  
 RT "Sequence and characterization of mxaB, a response regulator involved  
 in regulation of methanol oxidation, and of mxaW, a methanol-regulated  
 gene in Methylobacterium extorquens AM1."  
 RL FEMS Microbiol. Lett. 160:119-124(1998).  
 DR EMBL; AF017434; AAC46166.1; --  
 SQ SEQUENCE 129 AA; 13524 MW; 0AB4273418991F7C CRC64;

Query Match 76.5%; Score 13; DB 2; Length 129;  
 Best Local Similarity 50.0%; Pred. No. 1.9e+03;  
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7  
 |  
 |  
 |  
 DB 36 EAALL 41

RESULT 39  
 Q9YDZ3  
 ID Q9YDZ3 PRELIMINARY; PRT; 129 AA.  
 AC Q9YDZ3;  
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)  
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
 DE Hypothetical protein APE0776.  
 GN APE0776.

OS Aeropyrum pernix.  
 OC Archaea; Crenarchaeota; Thermoprotei; Desulfurococcales;  
 OC Desulfurococaceae; Aeropyrum.  
 OX NCBI\_TaxID=56636;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=K1;  
 RX MEDLINE=99310339; PubMed=10382966;  
 RA Kawarabayashi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,  
 RA Jin-no K., Takahashi M., Sekine M., Baba S.-I., Ankai A., Kosugi H.,  
 RA Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,  
 RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,  
 RA Yamazaki J., Kushiida N., Oguchi A., Aoki K.-I., Kubota K.,  
 RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.,  
 RT "Complete genome sequence of an aerobic hyper-thermophilic  
 crenarchaeon, Aeropyrum pernix K1."  
 RL DNA Res. 6:83-101(1999).  
 DR EMBL; AP000060; BAA79754.1; --  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 129 AA; 14730 MW; 0F67BB6C5E7478BA CRC64;

Query Match 76.5%; Score 13; DB 17; Length 129;  
 Best Local Similarity 50.0%; Pred. No. 1.9e+03;  
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7  
 |  
 |  
 |  
 DB 78 EAASLL 83

RESULT 40  
 Q9LYY0  
 ID Q9LYY0 PRELIMINARY; PRT; 136 AA.  
 AC Q9LYY0;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)  
 DE Hypothetical 15.6 kDa protein.  
 GN F15A17.80.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Bevan M., Terry N., Ardiles W., Buyssehaert C., Dasseville R.,  
 RA De Clerck R., De Keyser A., Neyt P., Rouze P., Van Den Daele H.,  
 RA Villarroel R., Gielens J., Van Montagu M., Bancroft I., Mewes H.W.,  
 RA Rudd S., Lemcke K., Mayer K.F.X.;  
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA EU Arabidopsis sequencing project;  
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AL163002; CAB86072.1; --  
 KW Hypothetical protein.  
 SQ SEQUENCE 136 AA; 15532 MW; 4F5736A0C709E34E CRC64;

Query Match 76.5%; Score 13; DB 10; Length 136;  
 Best Local Similarity 50.0%; Pred. No. 2e+03;  
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EXXXLL 7  
 |  
 |  
 |  
 DB 117 ETTSL 122

Search completed: November 5, 2003, 14:17:51  
 Job time: 36 secs

Hayes  
09/16/2008 40

09/16/2008 40

FILE 'REGISTRY' ENTERED AT 15:04:19 ON 05 NOV 2003  
338332 S E[3.]LL/SQSP

FILE 'REGISTRY' ENTERED AT 15:19:22 ON 05 NOV 2003  
18109 S L1 AND SQL=<100

FILE 'HCAPLUS' ENTERED AT 15:22:16 ON 05 NOV 2003  
3630 S L2

35 S L3 AND CLOSTRID?

8 S L4 AND ?TOXIN?

L8 ANSWER 1 OF 8 HCAPLUS COPYRIGHT 2003 ACS on STN

ACCESSION NUMBER: 2003:590711 HCAPLUS

DOCUMENT NUMBER: 139:129339

TITLE: Fluorophore-labeled peptides and FRET assays for

**clostridial toxins**

INVENTOR(S): Steward, Lance E.; Fernandez-Salas, Ester; Aoki,

Kel Roger

PATENT ASSIGNEE(S):

USA

U.S. Pat. Appl. Publ., 69 pp.

CODEN: USXXCO

Patent

English

FAMILY ACC. NUM. COUNT: 1

PATENT INFORMATION:

PATENT NO. KIND DATE APPLICATION NO. DATE

US 2003143651 A1 20030731 US 2001-942098 20010828

PRIORITY APPL. INFO.:

US 2001-942098 20010828

The present invention provides **clostridial toxin**

substrates useful in assaying for the protease activity of any

**clostridial toxin**, including botulinum

toxins of all serotypes as well as tetanus toxins.

A **clostridial toxin** substrate of the invention

contains a donor fluorophore; an acceptor having an absorbance

spectrum overlapping the emission spectrum of the donor fluorophore;

and a **clostridial toxin** recognition sequence

that includes a cleavage site, where the cleavage site intervenes

between the donor fluorophore and the acceptor and where, under the

appropriate conditions, resonance energy transfer is exhibited

between the donor fluorophore and the acceptor.

IT 566212-87-1

RT: PRP (Properties)

(unclaimed protein sequence; fluorophore-labeled peptides and

FRET assays for **clostridial toxins**)

L8 ANSWER 2 OF 8 HCAPLUS COPYRIGHT 2003 ACS on STN

ACCESSION NUMBER: 2003:356176 HCAPLUS

DOCUMENT NUMBER: 138:348758

TITLE: Endothelial-cell binding peptides for diagnosis

INVENTOR(S):

Gyuris, Jeno; Lamphere, Lou; Morris, Aaron J.;

Tsaioun, Katherine

GPC Biotech Inc., USA

SOURCE: PCT Int. Appl., 126 pp.

Searcher :

Shears

308-4994



CODEN: PIXXD2

Patent

English

FAMILY ACC. NUM. COUNT: 1

PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 2003037172	A2	20030508	WO 2002-US35258	20021101

W: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, GR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, ST, TM, TN, TR, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM

RW: GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, SK, TR, BF, BJ, CF, CG, CI, CM, GN, GQ, GW, ML, MR, NE, SN, TD, TG

US 2003166004 A1 20030904 US 2002-286457 20021101  
 PRIORITY APPLN. INFO.: US 2001-334822P P 20011101

AB The present invention relates to peptides and their derivs. which bind to endothelial cells and inhibit their proliferation in vitro assays, e.g., also referred to herein as endothelial cell binding peptide (ECBP) or ECBP sequence. These compns. may be combined with a pharmaceutically acceptable excipient or carrier and used to inhibit angiogenesis and angiogenesis-related diseases such as cancer, arthritis, macular degeneration, and diabetic retinopathy.

IT 518997-99-4 518999-26-3  
 RL: DGN (Diagnostic use); PRP (Properties); THU (Therapeutic use); BIOL (Biological study); USES (Uses)  
 (endothelial-cell binding peptides for diagnosis and therapy of angiogenesis-related disorders)

L8 ANSWER 3 OF 8 HCAPLUS COPYRIGHT 2003 ACS on STN

ACCESSION NUMBER: 2003:202825 HCAPLUS

DOCUMENT NUMBER: 138:233337

TITLE: FRET protease assays for botulinum serotype A/E

INVENTOR(S): Steward, Lance E.; Fernandez-Salas, Ester; Aoki, Kei Roger

PATENT ASSIGNEE(S): Allergan, Inc., USA

SOURCE: PCT Int. Appl., 168 pp.

CODEN: PIXXD2

DOCUMENT TYPE: Patent

LANGUAGE: English

FAMILY ACC. NUM. COUNT: 1

PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 2003020948	A2	20030313	WO 2002-US27145	20020822

W: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, GR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ,

Searcher : Shears 308-4994

09/620840

LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZM, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM  
BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, SK, TR, BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG  
US 2003143650 A1 20030731 US 2001-942024 20010828  
PRIORITY APPLN. INFO. :  
US 2001-942024 A 20010828

AB The present invention provides **clostridial toxin** substrates useful in assaying for the protease activity of botulinum serotype A/E toxins. A **clostridial toxin** substrate of the invention contains a donor fluorophore; an acceptor having an absorbance spectrum overlapping the emission spectrum of the donor fluorophore; and a **clostridial toxin** recognition sequence that includes a cleavage site, where the cleavage site intervenes between the donor fluorophore and the acceptor and where, under the appropriate conditions, resonance energy transfer is exhibited between the donor fluorophore and the acceptor.  
IT 501506-02-1  
RL: RFP (Properties)  
(unclaimed protein sequence; FRET protease assays for botulinum serotype A/E toxins)

L8 ANSWER 4 OF 8 HCAPLUS COPYRIGHT 2003 ACS on STN  
ACCESSION NUMBER: 2003:97984 HCAPLUS  
DOCUMENT NUMBER: 138:147749  
TITLE: Leucine-based motif and **clostridial neurotoxins**  
INVENTOR(S) : Steward, Lance E.; Fernandez-Salas, Ester; Herrington, Todd M.; Aoki, Kei Roger  
PATENT ASSIGNEE(S) : Allergan Sales, Inc., USA  
SOURCE: U.S. Pat. Appl. Publ., 33 pp., Cont.-in-part of U.S. Ser. No. 620,840.  
CODEN: USXXCO  
DOCUMENT TYPE: Patent  
LANGUAGE: English  
FAMILY ACC. NUM. COUNT: 2  
PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
US 2003027752	A1	20030206	US 2001-910346	20010720
PRIORITY APPLN. INFO. : US 2000-620840 A2 200000721				

AB Modified **neurotoxin** comprising **neurotoxin** including structural modification, wherein the structural modification alters the biol. persistence, such as the biol. half-life and/or a biol. activity of the modified **neurotoxin** relative to an identical **neurotoxin** without the structural modification. In one embodiment, methods of making the modified **neurotoxin** include using recombinant techniques. In another embodiment, methods of using the modified **neurotoxin** to treat conditions include treating various disorders, neuromuscular ailments and pain.  
IT 494824-08-7  
RL: RFP (Properties)

(unclaimed protein sequence; leucine-based motif and Clostridial neurotoxins)  
 Clostridial neurotoxins  
 392335-74-9 394737-43-0 394737-45-2  
 RT: PRP (Properties)  
 (unclaimed sequence; leucine-based motif and Clostridial neurotoxins)

L8 ANSWER 5 OF 8 HCAPLUS COPYRIGHT 2003 ACS on STN  
 ACCESSION NUMBER: 2002:121143 HCAPLUS  
 DOCUMENT NUMBER: 136:145953

TITLE: Complete genome sequence of Clostridium perfringens, an anaerobic flesh-eater

AUTHOR(S): Shimizu, Tohru; Ohtani, Kaori; Hirakawa, Hideki; Ohshima, Kenshiro; Yamashita, Atsushi; Shiba, Tadayoshi; Ogasawara, Naotake; Hattori, Masahira; Kuhara, Satoru; Hayashi, Hideo  
 DEPARTMENT OF MICROBIOLOGY, INSTITUTE OF BASIC MEDICAL SCIENCES, UNIVERSITY OF TSUKUBA, TSUKUBA, 305-8575, JAPAN  
 SOURCE: Proceedings of the National Academy of Sciences of the United States of America (2002), 99(2), 996-1001  
 CODEN: PNAS6; ISSN: 0027-8424  
 PUBLISHER: National Academy of Sciences  
 DOCUMENT TYPE: Journal  
 LANGUAGE: English

AB Clostridium perfringens is a Gram-pos. anaerobic

spore-forming bacterium that causes life-threatening gas gangrene and mild enterotoxemia in humans, although it colonizes as normal intestinal flora of humans and animals. The organism is known to produce a variety of toxins and enzymes that are responsible for the severe myonecrotic lesions. The complete 3,031,430-bp sequence of C. perfringens strain 13 comprises 2660 protein-coding regions and 10 rRNA genes, showing pronounced low overall G + C content (28.6%). The genome contains typical anaerobic fermentation enzymes leading to gas production but no enzymes for the tricarboxylic acid cycle or respiratory chain. Various saccharolytic enzymes were found, but many enzymes for amino acid biosynthesis were lacking in the genome. Twenty genes were newly identified as putative virulence factors of C. perfringens, and a total of five hyaluronidase genes were found that will also contribute to virulence. The genome anal. also proved an efficient method for finding four members of the two-component VirR/VirS regulon that coordinately regulates the pathogenicity of C. perfringens. Clearly, C. perfringens obtains various essential materials from the host by producing several degradative enzymes and toxins, resulting in massive destruction of the host tissues. The sequences are deposited in GenBank under Accession Nos. BA000016 (circular chromosome) and AF003515 (plasmid pCPF13).  
 393602-07-8 393602-38-5 393602-50-1  
 393602-67-0 393604-59-6 393604-89-2  
 393613-60-0 393618-67-2 393618-79-6  
 RT: BSU (Biological study, unclassified); PRP (Properties); BIOL (Biological study)  
 (amino acid sequence; complete genome sequence of Clostridium perfringens)  
 REFERENCE COUNT: 49  
 THERE ARE 49 CITED REFERENCES AVAILABLE  
 FOR THIS RECORD. ALL CITATIONS AVAILABLE

09/620840

IN THE RE FORMAT

L8 ANSWER 6 OF 8 HCAPLUS COPYRIGHT 2003 ACS on STN

ACCESSION NUMBER: 2002:90086 HCAPLUS

DOCUMENT NUMBER: 136:156405

TITLE:

Method for structural modifying  
Clostridial neurotoxins for

altering biological activity or persistence by  
leucine-based motifs

INVENTOR(S):

Steward, Lance E.; Fernandez-Salas, Ester;  
Herrington, Todd M.; Aoki, Kei Roger

PATENT ASSIGNEE(S):

Allergan Sales, Inc., USA  
PCT Int. Appl., 102 pp.

SOURCE:

DOCUMENT TYPE:

Patent

LANGUAGE:

English

FAMILY ACC. NUM. COUNT: 2

PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
------------	------	------	-----------------	------

WO 2002008268	A2	20020131	WO 2001-US23122	20010720
WO 2002008268	A3	20030220		

W:

AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, GU, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM  
RW: GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW, AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR, BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG

EP 1309618 A2 20030514 EP 2001-959115 20010720  
R: AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE, MC, PT, IE, SI, LT, LV, FI, RO, MK, CY, AL, TR  
BR 2001012715 A 20030520  
US 2000-620840 A 20000721  
WO 2001-US23122 W 20010720

PRIORITY APPLN. INFO.:

AB The invention provides a method for structural modifying botulinum toxin with leucine-based motifs. Modified neurotoxin comprising neurotoxin including structural modification, wherein the structural modification alters activity of the modified neurotoxin relative to an identical neurotoxin without the structural modification. In one embodiment, methods of making the modified neurotoxin include using recombinant techniques. In another embodiment, methods of using the modified neurotoxin to treat conditions include treating various disorders, neuromuscular ailments and pain.

IT

394746-46-4

RL: RRP (Properties)

(unclaimed protein sequence; method for structural modifying Clostridial neurotoxins for altering biol. activity or persistence by leucine-based motifs)

IT

392335-74-9 394737-43-0 394737-45-2

Searcher :

Shears

308-4994

03/620840

394737-51-0 394746-54-4 394746-55-5

RL: FRP (Properties)

(unclaimed sequence; method for structural modifying Clostridia neurotoxins for altering biol. activity or persistence by leucine-based motifs)

L8 ANSWER 7 OF 8 HCAPLUS COPYRIGHT 2003 ACS on STN

ACCESSION NUMBER: 2001:449905 HCAPLUS

DOCUMENT NUMBER: 135:60160

TITLE:

Immunogenic compositions against Helicobacter infection, polypeptides for use in the compositions and nucleic acid sequences encoding said polypeptides

INVENTOR(S):

PATENT ASSIGNEE(S):

SOURCE:

DOCUMENT TYPE:

LANGUAGE:

FAMILY ACC. NUM. COUNT: 4

PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
------------	------	------	-----------------	------

US 6248330	B1	20010619	US 1995-432697	19950502
------------	----	----------	----------------	----------

WO 9514093	A1	19950526	WO 1993-EP3259	19931119
------------	----	----------	----------------	----------

WO 9426901	A1	19941124	WO 1994-EP1625	19940519
------------	----	----------	----------------	----------

US 5843460	A	19981201	US 1995-467822	19950606
------------	---	----------	----------------	----------

US 6258359	B1	20010710	US 1995-466248	19950606
------------	----	----------	----------------	----------

WO 9634624	A1	19961107	WO 1996-EP1834	19960502
------------	----	----------	----------------	----------

AT, AU, AM, AZ, BB, BG, BR, BY, CA, CH, CN, CZ, DE, DK, EE, ES, FI, GB, GE, HU, IS, JP, KE, KG, KP, KR, KZ, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI				
--	--	--	--	--

RW: KE, LS, MW, SD, SZ, UG, AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, SF, BJ, CF, CG, CI, CM, GA, GN				
--	--	--	--	--

AU 9656934	A1	19961121	AU 1996-56934	19960502
------------	----	----------	---------------	----------

PRIORITY APPLN. INFO.:

WO 1993-EP3259	A	19930519	EP 1993-401309	A	19930519
----------------	---	----------	----------------	---	----------

WO 1993-EP3259	W	19931119	WO 1994-EP1625	A2	19940519
----------------	---	----------	----------------	----	----------

US 1995-432697	A2	19950502	US 1995-447177	A1	19950519
----------------	----	----------	----------------	----	----------

WO 1996-EP1834	W	19960502	US 1996-EP1834	W	19960502
----------------	---	----------	----------------	---	----------

AB

There is provided an immunogenic composition capable of inducing protective antibodies against Helicobacter infection characterized in that it comprises: i) at least one sub-unit of a urease structural polypeptide from Helicobacter pylori, or a fragment thereof, said fragment being recognized by antibodies reacting with Helicobacter fells urease, and/or at least one sub-unit of a urease structural polypeptide from Helicobacter fells, or a fragment thereof, said fragment being recognized by antibodies reacting with Helicobacter pylori urease; ii) and/or, a heat shock protein (Hsp),

Searcher : Shears 308-4994

or chaperonin, from Helicobacter, or a fragment of said protein.  
The preparation, by recombinant means, of such immunogenic comps. is  
also provided.  
127314-53-8, Urease (Proteus mirabilis clone PMID1003  
γ-subunit protein moiety reduced)  
RT: RRP (Properties)  
(unclamed protein sequence; immunogenic comps. against  
Helicobacter infection, polypeptides for use in the comps. and  
nucleic acid sequences encoding said polypeptides)  
REFERENCE COUNT: 62  
THERE ARE 62 CITED REFERENCES AVAILABLE  
FOR THIS RECORD. ALL CITATIONS AVAILABLE  
IN THE RE FORMAT

LB ANSWER 8 OF 8 HCAPLUS COPYRIGHT 2003 ACS on STN  
 ACCESSION NUMBER: 1999:614249 HCAPLUS  
 DOCUMENT NUMBER: 131:252536  
 TITLE: Assays for screening compounds which interact  
 with cation channel proteins, mutant prokaryotic  
 cation channel proteins, and uses thereof  
 INVENTOR(S): Mackinnon, Roderick  
 PATENT ASSIGNEE(S): The Rockefeller University, USA  
 SOURCE: PCT Int. Appl., 165 pp.  
 CODEN: PIXXD2  
 Patent  
 DOCUMENT TYPE: English  
 FAMILY ACC. NUM. COUNT: 2  
 PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 9947923	A2	19990923	WO 1999-US6307	19990322
WO 9947923	A3	20021003		

W:		RW:		PRIORITY APPLN. INFO.:	
AE, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GU, HU, ID, IL, IN, IS, JP, KE, KR, KZ, LG, MD, MG, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZA, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM, UA, UG, US, UZ, VN, YU, ZA, ZW, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG	CA 2323725	AA	19990923	CA 1999-2323725	19990322
	AU 9931988	A1	19991011	AU 1999-31988	19990322
	EP 1062508	A1	20001227	EP 1999-914058	19990322
	R: AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE, MC, PT, IE, FI				
					A 19980320

AB	Assays for screening potential drugs or agents that can interact and potentially bind to cation channel proteins, and potentially have uses in treating conditions related to the function of cation channel proteins is provided, along with prokaryotic cation channel proteins mutated to mimic eukaryotic cation channels, which can then be used in assays of the present invention.	244772-84-7, PN: W09947923 FIG: 1 unclaimed protein 244772-91-6, PN: W09947923 FIG: 1 unclaimed protein	RT: RRP (Properties)
IT			

09/620840

(unclamed protein sequence; assays for screening comps. which interact with cation channel proteins, mutant prokaryotic cation channel proteins, and uses thereof)

E294 THROUGH E317 ASSIGNED

FILE 'REGISTRY' ENTERED AT 15:30:59 ON 05 NOV 2003

L9 24 SEA FILE=REGISTRY ABB=ON PLU=ON (392335-74-9/BI OR 394737-43-0/BI OR 394737-45-2/BI OR 127314-53-8/BI OR 244772-84-7/BI OR 244772-91-6/BI OR 393602-07-8/BI OR 393602-38-5/BI OR 393602-50-1/BI OR 393602-67-0/BI OR 393604-59-6/BI OR 393604-89-2/BI OR 393613-60-0/BI OR 393618-67-2/BI OR 393618-79-6/BI OR 394737-51-0/BI OR 394746-46-4/BI OR 394746-54-4/BI OR 394746-55-5/BI OR 494824-08-7/BI OR 501506-02-1/BI OR 518997-99-4/BI OR 518999-26-3/BI OR 566212-87-1/BI)

L10 24 L9 AND L1

L10 ANSWER 1 OF 24 REGISTRY COPYRIGHT 2003 ACS on STN  
RN 566212-87-1 REGISTRY  
CN L-Methionine, L- $\alpha$ -aspartyl-L-leucyl-L-valyl-L-alanyl-L-glutamyl-L-arginylglycyl-L- $\alpha$ -glutamyl-L-arginyl-L-leucyl-L- $\alpha$ -glutamyl-L- $\alpha$ -aspartyl-L-lysyl-L-threonyl-L- $\alpha$ -glutamyl-L-asparaginyl-L-leucyl-L-valyl-L-lysyl-L-threonyl-L-seryl-L-threonyl-L-arginyl-L-leucyl-L-phenylalanyl-L-aspartyl-L-seryl-L-valyl-L-threonyl-L-phenylalanyl-L-lysyl-L-threonyl-L-threonyl-L-arginyl-L-methionyl-L-cysteinyl- (9CI) (CA)  
OTHER NAMES:  
CN 73: PN: US20030143651 SEQID: 62 unclamed protein  
CI MAN  
SQL 40

SEQ 1 DLVAQRGERL ELLIDKTENTL VDSSVTFKTT SRNLARACM  
HITS AT: 8-13  
\*\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

REFERENCE 1: 139:129339

L10 ANSWER 2 OF 24 REGISTRY COPYRIGHT 2003 ACS on STN  
RN 518999-26-3 REGISTRY  
CN L-Alanine, L-seryl-L-seryl-L-asparaginyl-L-seryl-L- $\alpha$ -glutamyl-L-leucyl-L-tryptophyl-L-proyl-L-leucyl-L-leucyl-L-seryl- (9CI) (CA)  
SQL 12  
INDEX NAME)

SEQ 1 SSNSLWPL SA  
HITS AT: 5-10  
=====

REFERENCE 1: 138:348758

L10 ANSWER 3 OF 24 REGISTRY COPYRIGHT 2003 ACS on STN  
RN 518997-99-4 REGISTRY

Searcher : Shears 308-4994

09/620840

CN L-Isoleucine, L-valyl-L-alanyl-L-lysyl-L-histidyl-L- $\alpha$ -glutamylglycyl-L-valyl-L-seryl-L-leucyl-L-leucyl-L-arginyl- (9CI) (CA INDEX NAME)  
SOL 12  
SEQ 1 VAKHEGVSL RI  
HITS AT: 5-10  
=====

REFERENCE 1: 138:348758

L10 ANSWER 4 OF 24 REGISTRY COPYRIGHT 2003 ACS on STN

RN 501506-02-1 REGISTRY

CN L-methionine, L- $\alpha$ -aspartyl-L-leucyl-L-valyl-L-alanyl-L-glutamyl-L-arginylglycyl-L- $\alpha$ -glutamyl-L-arginyl-L-leucyl-L- $\alpha$ -glutamyl-L-leucyl-L- $\alpha$ -aspartyl-L-lysyl-L-threonyl-L- $\alpha$ -glutamyl-L-leucyl-L-valyl-L-phenylalanyl-L- $\alpha$ -aspartyl-L-seryl-L-seryl-L-valyl-L-threonyl-L-lysyl-L-threonyl-L-threonyl-L-seryl-L-arginyl-L-asparagyl-L-leucyl-L-alanyl-L-arginyl-L-alanyl-L-methionyl-L-cysteinyl- (9CI) (CA INDEX NAME)  
OTHER NAMES:  
CN 62: PN: W003020948 SEQID: 62 unclamed protein  
CI MAN  
SOL 40  
SEQ 1 DLVAQRGERL ELLIDKTEML VDSSVTFKTT SRNLARAMCM  
HITS AT: 8-13  
=====

\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

REFERENCE 1: 138:233337

L10 ANSWER 5 OF 24 REGISTRY COPYRIGHT 2003 ACS on STN

RN 494824-08-7 REGISTRY

CN Peptide, (Xaa-glu-Xaa-Xaa-Leu-Leu) (9CI) (CA INDEX NAME)  
OTHER NAMES:  
CN 2: PN: US20030027752 SEQID: 2 unclamed protein  
CI MAN  
SOL 7  
SEQ 1 XEXXXLL  
HITS AT: 2-7  
=====

\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

REFERENCE 1: 138:147749

L10 ANSWER 6 OF 24 REGISTRY COPYRIGHT 2003 ACS on STN

RN 394746-55-5 REGISTRY

CN 26: PN: W00208268 FIGURE: 10 unclamed sequence (9CI) (CA INDEX NAME)  
CI MAN  
SOL 87  
SEQ 1 MMHHHHHHSS GLVFRSGMK ETAAAKFERQ HMDSPLDGLTD DDDKAMGSFV

Searcher : Shears 308-4994



09/620840

51 NKQFNKDPV NFKLKNFTG LKEFYKLLCV RGIITSK

HITS AT: 73-78

REFERENCE 1: 136:156405

L10 ANSWER 7 OF 24 REGISTRY COPYRIGHT 2003 ACS on STN

RN 394746-54-4 REGISTRY

CN 25: PN: W00208268 FIGURE: 10 unclaimed sequence (9CI) (CA INDEX NAME)

CI MAN

SQ 77

SEQ 1 MHHHHHSSG LVPKSGMKE TAAAKFERQM HDSPDLGTDD DDKAMYKDPV

51 NFKLKNFTG LKEFYKLLCV RGIITSK

HITS AT: 63-68

REFERENCE 1: 136:156405

L10 ANSWER 8 OF 24 REGISTRY COPYRIGHT 2003 ACS on STN

RN 394746-46-4 REGISTRY

CN Peptide, (Xaa-Glu-Xaa-Xaa-Leu-Leu) (9CI) (CA INDEX NAME)

OTHER NAMES:

CN 3: PN: W00208268 SEQID: 2 unclaimed protein

CI MAN

SQ 7

SEQ 1 XEXXXLL

HITS AT: 2-7

\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

REFERENCE 1: 136:156405

L10 ANSWER 9 OF 24 REGISTRY COPYRIGHT 2003 ACS on STN

RN 394737-51-0 REGISTRY

CN L-Leucine, L-asparaginyl-L- $\alpha$ -glutamyl-L-seryl-L-prolyl-L-leucyl- (9CI) (CA INDEX NAME)

OTHER NAMES:

CN 21: PN: W00208268 SEQID: 21 unclaimed sequence

SQ 7

SEQ 1 NEQSPLT

HITS AT: 2-7

REFERENCE 1: 136:156405

L10 ANSWER 10 OF 24 REGISTRY COPYRIGHT 2003 ACS on STN

RN 394737-45-2 REGISTRY

CN L-Leucine, L-alanyl-L- $\alpha$ -glutamyl-L-valyl-L-glutamyl-L-alanyl-L-leucyl- (9CI) (CA INDEX NAME)

OTHER NAMES:

CN 12: PN: US20030027752 SEQID: 12 unclaimed sequence

CN 12: PN: W00208268 SEQID: 12 unclaimed sequence

SQ 7

Searcher : Shears 308-4994

SEQ 1 AEVQALL  
=====

HITS AT: 2-7

REFERENCE 1: 138:147749  
REFERENCE 2: 136:156405

L10 ANSWER 11 OF 24 REGISTRY COPYRIGHT 2003 ACS on STN  
RN 394737-43-0 REGISTRY  
CN L-Leucine, L-seryl-L- $\alpha$ -glutamyl-L-arginyl-L-aspartyl-L-valyl-L-leucyl- (9CI) (CA INDEX NAME)  
OTHER NAMES:  
CN 10: PN: US20030027752 SEQID: 10 unclaimed sequence  
CN 10: PN: W00208268 SEQID: 10 unclaimed sequence  
SOL 7

SEQ 1 SERDVL

=====

HITS AT: 2-7

REFERENCE 1: 138:147749  
REFERENCE 2: 136:156405

L10 ANSWER 12 OF 24 REGISTRY COPYRIGHT 2003 ACS on STN  
RN 393618-79-6 REGISTRY  
CN Protein (Clostridium perfringens strain 13 gene CPE1618) (9CI) (CA INDEX NAME)  
OTHER NAMES:  
CN GenBank BAB81324  
CN GenBank BAB81324 (Translated from: GenBank AP003191)  
SOL 85  
CI MAN

SEQ 1 MEKQIMAFR DEGNKVEFE VVAKIYLGEK NKKYIVLSP VEGNGDEADD  
51 EVFRVDKVD SVEYNLVEDD EEFRLVKKKEY KKLTY

=====

HITS AT: 79-84

REFERENCE 1: 136:145953

L10 ANSWER 13 OF 24 REGISTRY COPYRIGHT 2003 ACS on STN  
RN 393618-67-2 REGISTRY  
CN Protein (Clostridium perfringens strain 13 gene CPE1606) (9CI) (CA INDEX NAME)  
OTHER NAMES:  
CN GenBank BAB81312  
CN GenBank BAB81312 (Translated from: GenBank AP003191)  
SOL 99  
CI MAN

SEQ 1 MGFLNNDEG NDFLLADTYS KDDIFWITLR GLISKGKIDE AEDMLFKEAF  
=====

HITS AT: 10-15

REFERENCE 1: 136:145953

L10 ANSWER 14 OF 24 REGISTRY COPYRIGHT 2003 ACS on STN

RN 393613-60-0 REGISTRY

CN Protein (Clostridium perfringens strain 13 gene CPE1101) (9CI) (CA)

OTHER NAMES:

CN GenBank BAB80807

CN GenBank BAB80807 (Translated from: GenBank AP003189)

CI MAN

SOL 77

SEQ 1 MSVIREDFRR LSKQIMLED LQIIKEITGT SKITNDTVQD LFLALQSHLL

51 LVEEGKTEIN TEFTEITSL VAKLYNR

HITS AT: 65-70

REFERENCE 1: 136:145953

L10 ANSWER 15 OF 24 REGISTRY COPYRIGHT 2003 ACS on STN

RN 393604-89-2 REGISTRY

CN Protein (Clostridium perfringens strain 13 gene CPE0230) (9CI) (CA)

OTHER NAMES:

CN GenBank BAB79936

CN GenBank BAB79936 (Translated from: GenBank AP003186)

CI MAN

SOL 61

SEQ 1 MGNIKDRIIV CRCKVSEGT IIEAIKNGAD TYEKVKKETG ANTYGCFACR

51 LEIKKLEEN K

HITS AT: 52-57

REFERENCE 1: 136:145953

L10 ANSWER 16 OF 24 REGISTRY COPYRIGHT 2003 ACS on STN

RN 393604-59-6 REGISTRY

CN Protein (Clostridium perfringens strain 13 gene CPE0200) (9CI) (CA)

OTHER NAMES:

CN GenBank BAB79906

CN GenBank BAB79906 (Translated from: GenBank AP003185)

CI MAN

SOL 53

SEQ 1 MFTNETKDLT AGLYQKYGCT QEVLTLSNII DKIIIVKEQSE RLKEYYKSRK

HITS AT: 51 NNI

HITS AT: 5-10

REFERENCE 1: 136:145953

L10 ANSWER 17 OF 24 REGISTRY COPYRIGHT 2003 ACS on STN

RN 393602-67-0 REGISTRY

CN Protein (Clostridium perfringens strain 13 gene CPE0008) (9CI) (CA)

OTHER NAMES:

INDEX NAME)

OTHER NAMES:

Searcher :

Shears

308-4994

09/620840

CN GenBank BAB79714  
CN GenBank BAB79714 (Translated from: GenBank AP003185)  
CI MAN  
SOL 83

SEQ 1 MSTCIINSQG VAIRINVSADI SVTSRSAMGV KIMRTLEDEA VTIAKISGD  
51 TNEEEETITL LDESEDNTNII VEEVAITEVKD EEN  
=====

HITS AT: 56-61

REFERENCE 1: 136:145953

L10 ANSWER 18 OF 24 REGISTRY COPYRIGHT 2003 ACS on STN

RN 393602-50-1 REGISTRY  
CN Protein (Clostridium perfringens strain 13 plasmid PCP54) gene PCP54)  
OTHER NAMES: (9CI) (CA INDEX NAME)

CN GenBank BAB62492  
CN GenBank BAB62492 (Translated from: GenBank AP003515)  
CI MAN

SOL 64

SEQ 1 MARGNNRNNNV KNPIKETVSM ILESKGISYD TWLEEAHMKL IFENLELNE  
=====

51 GLALKKELEG NSGE

HITS AT: 43-48

REFERENCE 1: 136:145953

L10 ANSWER 19 OF 24 REGISTRY COPYRIGHT 2003 ACS on STN

RN 393602-38-5 REGISTRY  
CN Protein (Clostridium perfringens strain 13 plasmid PCP13 gene PCP42)  
OTHER NAMES: (9CI) (CA INDEX NAME)

CN GenBank BAB62480  
CN GenBank BAB62480 (Translated from: GenBank AP003515)  
CI MAN

SOL 67

SEQ 1 MPSTAPCVIA ISPYALKVSI SSSYQFTIKY TITINVEFNE FEKVNLTESI  
=====

51 GALKKKFQYFH FLKGVVI

HITS AT: 42-47

REFERENCE 1: 136:145953

L10 ANSWER 20 OF 24 REGISTRY COPYRIGHT 2003 ACS on STN

RN 393602-07-8 REGISTRY  
CN Transposase (Clostridium perfringens strain 13 plasmid PCP13 gene  
OTHER NAMES: PCP11) (9CI) (CA INDEX NAME)

CN GenBank BAB62449

CN GenBank BAB62449 (Translated from: GenBank AP003515)

CI MAN

SOL 79

SEQ 1 MKMLYLSREK VNKKWTRNYP NMDLVINEELK ILLNEYLISKE CKKKRALRKOC  
=====

Searcher : Shears 308-4994

09/620840

51 SKSSKIQT KC ILTPVGN I HNKFYSRST  
HITS AT: 28-33

REFERENCE 1: 136:145953

L10 ANSWER 21 OF 24 REGISTRY COPYRIGHT 2003 ACS on STN

RN 392335-74-9 REGISTRY

CN L-Leucine, L-phenylalanyl-L- $\alpha$ -glutamyl-L-phenylalanyl-L-tyrosyl-L-lysyl-L-leucyl- (9CI) (CA INDEX NAME)

OTHER NAMES:

CN 7: PN: US20030027752 SEQID: 7 unclaimed sequence  
CN 7: PN: W00208268 SEQID: 7 unclaimed sequence

SOL 7

SEQ 1 FEYKLT

HITS AT: 2-7

REFERENCE 1: 138:147749

REFERENCE 2: 136:156405

L10 ANSWER 22 OF 24 REGISTRY COPYRIGHT 2003 ACS on STN

RN 244772-91-6 REGISTRY

CN PN: W09947923 FIG: 1 unclaimed protein (9CI) (CA INDEX NAME)

SOL 60

SEQ 1 GFVSAPLF SI ETEYXXXXXY RVITDKCP EG ILLTIQSVL GSI VNAFMVG

51 CMFVKISQPK

HITS AT: 29-34

REFERENCE 1: 131:252536

L10 ANSWER 23 OF 24 REGISTRY COPYRIGHT 2003 ACS on STN

RN 244772-84-7 REGISTRY

CN PN: W09947923 FIG: 1 unclaimed protein (9CI) (CA INDEX NAME)

SOL 58

SEQ 1 TYWECVYLIM VTMSXXXXXD VYAKTTLGRL FMVFILGL AMFASVYPEI

51 IELIGNRK

HITS AT: 4-9

REFERENCE 1: 131:252536

L10 ANSWER 24 OF 24 REGISTRY COPYRIGHT 2003 ACS on STN

RN 127314-53-8 REGISTRY

CN Urease (Proteus mirabilis clone PMID1003  $\gamma$ -subunit protein moiety reduced) (9CI) (CA INDEX NAME)

OTHER NAMES:

CN 18: PN: US6248330 SEQID: 23 unclaimed protein

SOL 100

SEQ 1 MELTPREKDK LTLFTAGLVA ERLAKGLKL NYPERVALIS CAIMEGAREG

Searcher : Shears 308-4994

09/620840

51 KTVAQLMSEG RTVLTAQVM EGVPEMIKDV QVECTFPDGT KLVSIHSPIV  
HITS AT: 7-12  
=====

\*\*RELATED SEQUENCES AVAILABLE WITH SEOLINK\*\*

REFERENCE 1 : 135:60160

REFERENCE 2 : 112:231815

FILE 'HOME' ENTERED AT 15:31:20 ON 05 NOV 2003

Searcher : Shears 308-4994